UNIV. OF
TORONTO

TRANSACTIONS

OF THE

AMERICAN MATHEMATICAL SOCIETY

EDITED BY

GEORGE DAVID BIRKHOFF

ARTHUR BYRON COBLE

LUTHER PFAHLER EISENHART

WITH THE COÖPERATION OF

OLIVER E. GLENN	WALLIE A. HURWITZ	DUNHAM JACKSON
EDWARD KASNER	AUBREY J. KEMPNER	WILLIAM R. LONGLEY
HOWARD H. MITCHELL	CHARLES N. MOORE	ROBERT L. MOORE
FOREST R. MOULTON	FRANCIS R. SHARPE	J. H. M. WEDDERBURN
	ERNEST J. WILCZYNSKI	

VOLUME 24

1922

201280
12.3.26

PUBLISHED BY THE SOCIETY
LANCASTER, PA., AND NEW YORK
1923-24

QA
1
A53
V. 24

LANCASTER PRESS, INC.
LANCASTER, PA.

TABLE OF CONTENTS

VOLUME 24, 1922

PAGE

BELL, E. T., of Seattle, Wash. Anharmonic polynomial generalizations of the numbers of Bernoulli and Euler 89

BLISS, G. A., of Chicago, Ill. Birational transformations simplifying singularities of algebraic curves 274

BLUMBERG, H., of Urbana, Ill. New properties of all real functions 113

COBLE, A. B., of Urbana, Ill. Associated sets of points 1

COOLIDGE, J. L., of Cambridge, Mass. The Gaussian law of error for any number of variables 135

CURTISS, D. R., of Evanston, Ill. A note on the preceding paper (by J. L. WALSH) 181

EVERETT, H. S., of Lewisburg, Pa. Determination of all general homogeneous polynomials expressible as determinants whose elements are homogeneous polynomials 185

GRAUSTEIN, W. C., of Cambridge, Mass. Representations of a complex point by pairs of ordered real points 245

HAZLETT, OLIVE C., of South Hadley, Mass. A symbolic theory of formal modular covariants 286

MILLER, G. A., of Urbana, Ill. I-conjugate operators of an abelian group 70

MOORE, C. N., of Cincinnati, Ohio. Generalized limits in general analysis, first paper 79

MULLIKIN, ANNA M., of Austin, Tex. Certain theorems relating to plane connected point sets 144

PÓLYA, G., of Zurich, Switzerland. On the mean-value theorem corresponding to a given linear homogeneous differential equation 312

RITT, J. F., of New York, N. Y. On algebraic functions which can be expressed in terms of radicals 21, 324

SHAW, J. B., of Urbana, Ill. General vector calculus 195

TURNER, J. S., of Ames, Iowa. A fundamental system of invariants of a modular group of transformations 129

VAN VLECK, E. B., of Madison, Wis. Non-loxodromic substitutions and groups in n dimensions 255

WALSH, J. L., of Cambridge, Mass. On the location of the roots of the jacobian of two binary forms, and of the derivative of a rational function 31

——. On the location of the roots of certain types of polynomials 163

Errata, volume 24 324

ASSOCIATED SETS OF POINTS*

BY

ARTHUR B. COBLE†

INTRODUCTION

Two sets of n points ordered with respect to each other, the one, P_n^k, in a linear space S_k, determined by the equations

$$(up_1) = 0, \qquad (up_2) = 0, \qquad \cdots, \qquad (up_n) = 0,$$

and the other Q_n^{n-k-2}, in a linear space S_{n-k-2}, determined by the equations

$$(vq_1) = 0, \qquad (vq_2) = 0, \qquad \cdots, \qquad (vq_n) = 0,$$

are called *associated sets* if the factors of proportionality in the coördinates of the points can be so chosen that an identity in u, v exists of the following form:

$$(1) \qquad (up_1)(vq_1) + (up_2)(vq_2) + \cdots + (up_n)(vq_n) \equiv 0.$$

This relation, obviously mutual, between the two sets is such that either set uniquely defines the other to within projective modifications. Some general properties of such sets have been given by the writer.‡

A characteristic algebraic property of two associated sets is that complementary determinants formed from the matrices of the coördinates of the two sets of points when taken so that (1) is satisfied are proportional. A characteristic geometric property is the following: On $k+3$ of the points of P_n^k there is a unique rational norm curve N^k upon which the $k+3$ points determine a set of $k+3$ parameters; on the complementary set of $n-k-3$ points of Q_n^{n-k-2} there is a pencil of linear spaces S_{n-k-3} whose members on the remaining $k+3$ points determine a set of $k+3$ parameters; these two sets of $k+3$ parameters are projective.

Unless $k = n-k-2$ the associated sets are in spaces of different dimension. Conventional methods of passing from one space to another are the process of *mapping* the space of lower dimension upon that of higher dimension, and the process of *projecting* from the space of higher dimension upon the

* Read before the American Mathematical Society at Evanston, Dec. 29, 1922.

† This investigation has been pursued under the auspices of the Carnegie Institution of Washington, D. C.

‡ A. B. Coble, *Point sets and allied Cremona groups* (I), these Transactions, vol. 16 (1915), p. 155, in particular §§ 1, 2 and theorems (25), (26); also (II), vol. 17 (1916), p. 345, § 4 (16). These are cited as P. S. I or II.

one of lower dimension. Thus in the simple case of P_n^1, n points $x_0^{(i)}$, $x_1^{(i)}$ $(i = 1, \cdots, n)$ on a line, the line is mapped by means of the totality of binary $(n-3)$-ics in x, i.e., by $y_0 = (\alpha_0 x)^{n-3}$, $\cdots$, $y_{n-3} = (\alpha_{n-3} x)^{n-3}$, upon the points y of a rational norm curve N^{n-3} in S_{n-3} in such a way that P_n^1 is mapped upon its associated Q_n^{n-3}. On the other hand Q_n^{n-3} is projected from any S_{n-5} which is $(n-4)$-secant to N^{n-3} upon its associated P_n^1.

Two problems considered in this paper are: When $n - k - 2 \geqq k$ can the space S_k be mapped upon the space S_{n-k-2} so that the set P_n^k is mapped upon the set Q_n^{n-k-2}?; when $n - k - 2 > k$ can the set Q_n^{n-k-2} be projected upon the set P_n^k? For $k = 2$ the first problem is solved in § 1, the second in § 2. For $k = 3$ the first problem is solved in § 3. For the general set P_9^3 there appears to be no solution to the second problem and this probably would be true of further sets also.

In § 4 *particular* sets, i.e., those for which n, k have particular values, are considered. Each of these presents its own peculiarities. Also *special* sets, i.e., those which for given n, k satisfy in addition some projective conditions, receive some attention. Those conditions which are invariant under regular Cremona transformation of the set (cf. P. S. II, § 4) are especially emphasized. Their form in the two sets is often very diverse. Thus if P_9^3 is on a quadric with a node, then the associated Q_9^4 is on a rational quintic curve and conversely. In this section the discussion is carried through the values $n \leqq 10$.

The results obtained for the sets of nine and ten nodes of the rational sextic and of the symmetroid are useful in connection with the author's investigations of the modular functions of genus four attached to these figures.*

1. Mapping of P_n^2 upon its associated Q_n^{n-4}

The space S_2 is mapped upon the space S_{n-4} by means of a linear system Σ of ∞^{n-4} plane curves. The points of the plane are mapped upon a 2-way in S_{n-4} of order λ where λ is the number of variable intersections of two curves of Σ. The intersections of this 2-way by the linear S_{n-5}'s contained in S_{n-4} correspond in S_2 to the curves of Σ. We have therefore to find a system Σ so related to the set of points P_n^2 that the additional condition that three points of P_n^2 are on a line has as a consequence that there must exist a curve of Σ on the remaining $n-3$ points of P_n^2 and therefore also that the corresponding $n-3$ points of Q_n^{n-4} lie upon an S_{n-5} in S_{n-4}. This ensures the proportionality of complementary determinants in the matrices of the two point sets. Of course this requirement may not define the system Σ and we seek merely a simple system Σ with the required property.

* A part of this work appears in abstract in the Proceedings of the National Academy of Sciences, vol. 7 (1921), (I) p. 245; (II) p. 334. These are cited as Proc. I or II.

The cases where n is even and n is odd are slightly different and we begin with the mapping of a P^2_{2j+3} upon its associated Q^{2j-1}_{2j+3}. In S_2 pass through P^2_{2j+3} a proper curve C of order j with a $(j-2)$-fold point at a point r. Also pass through P^2_{2j+3} a proper curve D of order $(j+1)$ with a $(j-1)$-fold point at r which meets C in $(2j-5)$ points $s_1, s_2, \cdots, s_{2j-5}$. If L_1, L_2 are distinct lines on r, then D, CL_1, CL_2 cut out the same set S upon C, so that the set S lies in an I^{2j-5}_{j-3} on C. The choice of the points r, s thus depends upon $2j-5$ constants when P^2_{2j+3} is given. Let A, B respectively be arbitrary sets of $(j-2)$ and $(j-3)$ lines on r. Then in $AC+BD=0$ we have a system of curves of order $(2j-2)$ with a $(2j-4)$-fold point at r, on P^2_{2j+3}, and on the points S. The parameters in A and B are essential. For if $AC+BD \equiv A'C+B'D$ then $(A-A')C \equiv (B'-B)D$, whence $A \equiv A'$ and $B \equiv B'$, since C and D are proper curves. Thus the system $AC+BD=0$ contains ∞^{2j-4} curves. If three points of P^2_{2j+3} are on a line, a curve of the system can be passed through $(2j-4)$ further points of the line, which therefore will contain the line as a factor. The complementary factor will be a curve of the required system Σ of order $(2j-3)$ with a $(2j-4)$-fold point at r and on the set S, and this curve will pass through the complementary set of $2j$ points of P^2_{2j+3}. Hence the system Σ will map the set P^2_{2j+3} upon its associated set.

For the case n even, or a P^2_{2j+2}, we pass through P^2_{2j+2} two proper curves C, D of order j with a common $(j-2)$-fold point r which meet again in $(2j-6)$ points S. Here the choice of r, s depends upon $2j-6$ constants. Let A, B be arbitrary sets of $(j-3)$ lines on r. Then in $AC+BD=0$ we have a linear system of dimension $(2j-5)$ of curves of order $(2j-3)$ with a $(2j-5)$-fold point at r, on P^2_{2j+2}, and on the set S. If three of the points of P^2_{2j+2} are on a line, one curve of the system contains this line as a factor, whence one curve of the required system Σ of order $(2j-4)$ with a $(2j-5)$-fold point at r and on S will pass through the complementary set of $(2j-1)$ points of P^2_{2j+2}. This system Σ therefore effects the required mapping. Hence

THEOREM 1. *The plane set of points P^2_n is mapped upon its associated Q^{n-4}_n by a linear system of curves of order $(n-6)$ with an $(n-7)$-fold point at r and on a set of $(n-8)$ points S in such a way that the plane is mapped upon the normal 2-way, M^{n-5}_2, of order $(n-5)$ in S_{n-4}. If n is even the points S are the further intersections of two proper curves of order $(n-2)/2$ with a common $(n-6)/2$-fold point at the arbitrarily chosen point r and on the given set P^2_n. If n is odd the points S are the further intersections of two proper curves of order $(n-3)/2$ and $(n-1)/2$ with respectively $(n-7)/2$- and $(n-5)/2$-fold points at r and on P^2_n. For given P^2_n the choice of the points r, S depends upon $(n-8)$ constants.*

The mapping described above becomes evanescent for $n = 6$ and $n = 7$. In the case of P_6^2 let a pencil of cubics on P_6^2 meet again in s_1, s_2, s_3. Then conics on S map P_6^2 upon its associated Q_6^2. For if three of the points of P_6^2 are on a line, the complementary three are on a conic with s_1, s_2, s_3 and therefore map into three points of Q_6^2 on a line. Hence

THEOREM 2. *Six corresponding point pairs of a quadratic transformation are associated P_6^2, Q_6^2 if P_6^2 and the singular triangle of the transformation are the base points of a pencil of cubics.*

In the case of P_7^2 we pass a pencil of cubics through P_7^2 to meet again in s_1, s_2. Then conics on s_1, s_2 map P_7^2 upon its associated Q_7^3 in S_3. In this mapping the plane becomes a quadric on Q_7^3 and the points on the line $\overline{s_1 s_2}$ become the directions on this quadric about the eighth base point of the net of quadrics on Q_7^3. Thus to the ∞^2 possible choices of the pair s_1, s_2 there correspond the set Q_7^3 and the ∞^2 quadrics on it.

We observe also that the cases $n = 8$, $n = 9$ are exceptional in that for $P_8^2 r$ is the ninth base point of the pencil of cubics on P_8^2 and that for $P_9^2 r$ is a point on the cubic determined by P_9^2. For further cases r may be taken in general position.

2. THE PROJECTION OF Q_{k+4}^k UPON ITS ASSOCIATED P_{k+4}^2

We now consider the set Q_{k+4}^k as given in S_k and ask for spaces L of dimension $k - 3$ such that under projection from L, the set Q will become its associated set in the plane. Two lemmas are needed.

LEMMA 1. *The S_{k-2} π determined by L and q_1 is a $(k-1)$-secant space of the norm curve N_1^k on $q_2, \cdots, q_{k+4}$.*

For if τ is the parameter of the pencil of S_{k-1}'s on π and t the parameter on N_1^k the incidence condition of S_{k-1} τ and point t is a $(1, k)$ relation on τ, t which in general would have only $k + 1$ pairs τ, t in common with any $(1, 1)$ relation on τ, t. If this $(1, 1)$ relation is the projectivity mentioned in the introduction between the parameter τ of the line pencil on p_1 in S_2 and the parameter t of N_1^k, then it is satisfied by the $k + 3$ pairs t, τ determined by $q_2, \cdots, q_{k+4}$. Therefore the projectivity determines a $(1, 1)$ relation which is a factor of the $(1, k)$ relation. The complementary factor of degree $k - 1$ in t determines the points of N_1^k on π. Thus the $k + 4$ norm curves on the sets of $k + 3$ points q selected from Q_{k+4}^k are projected from L into $k + 4$ rational k-ics in the plane on the points of P_{k+4}^2 and with respectively a $(k - 1)$-fold point at each point of P_{k+4}^2. This remark is utilized in Theorem 5.

LEMMA 2. *Quadrics on $q_2, \cdots, q_{k+4}$ cut π in quadrics apolar to a unique quadric Q_π in π and L in π is the polar S_{k-3} of q_1 as to Q_π.*

For the $\binom{k}{2}$ linearly independent quadrics on N_1^k cut π in $\binom{k}{2}$ sections on

the $k-1$ points common to π and N_1^k, whence of these only $\binom{k}{2}-(k-1)$ are linearly independent in π. Therefore $k-1$ quadrics on N_1^k contain π and the $\binom{k+2}{2}-(k+3)$ quadrics in S_k on $q_2, \cdots, q_{k+4}$ cut π in at most $\binom{k-2}{2}-(k+3)-(k-1)=\binom{k}{2}-1$ linearly independent quadrics all of which are apolar to at least one quadric Q_π in π. Moreover the S_2 on three points of $q_2, \cdots, q_{k+4}$ and the S_{k-1} on the remaining k points meet π respectively in a point and S_{k-3} which are pole and polar as to Q_π and thereby Q_π is uniquely determined. For any S_{k-1} on S_2 together with the given S_{k-1} constitute a quadric on $q_2, \cdots, q_{k+4}$ and meet π in a pair of S_{k-3}'s apolar to Q_π. Finally, if three points of $q_2, \cdots, q_{k+4}$ are in an S_{k-1} with L and therefore project from L into three points of a line in S_2, then the remaining k points and q_1 must be in an S_{k-1} which meets π in an S_{k-3} on q_1. Hence the point, S_{k-3} of π mentioned above are such that when the point is on L then the S_{k-3} must be on q_1, which requires that q_1, L be pole and polar as to Q_π.

In order to put all the points of the set Q on the same footing we now prove

THEOREM 3. *Given Q_{k+4}^k in S_k there exist ∞^{k-3} spaces L of dimension $k-3$ such that all the quadrics on L and any $k+3$ of the points Q meet again at the remaining point of Q, or also such that all the quadrics on the points Q and $\binom{k-1}{2}-1$ points of L contain L. From any one of these spaces L the set Q_{k+4} is projected into its associated P_{r+4}^2.*

For there are ∞^{k-1} S_{k-2}'s which are $(k-1)$-secant spaces of N_1^k each with ∞^{k-2} points, so that on q_1 there are ∞^{k-3} such spaces π. In any such space π choose L to be the polar S_{k-3} of q_1 as to the quadric Q_π determined as in Lemma 2. Then all the quadrics of S_k on $q_2, \cdots, q_{k+4}$ which contain L cut π in another S_{k-3} on q_1 and L has the first property described in the theorem. That all the S_{k-3}'s L of the theorem are found among the $(k-1)$-secant spaces π of N_1^k on q_1 is proved as follows. If, as given, quadrics on $q_2, \cdots, q_{k+4}$ and L meet again in q_1, then the $\binom{k+2}{2}-\binom{k-1}{2}-(k+3)=2k-3$ linearly independent quadrics of this sort meet $\pi\,[L, q_1]$ in a linear system of S_{k-3}'s on p_1 of which only $k-2$ are linearly independent in π. Hence $k-1$ of the quadrics contain the S_{k-2} π and therefore meet in a N_1^k (necessarily on $q_2, \cdots, q_{k+4}$) which is $(k-1)$-secant to π. We observe that the configuration Q_{k+4}^k, L is the generalization of the set of eight base points of a net of quadrics as one of the points is enlarged in dimension. To prove the last statement in the theorem we note that if $q_2, \cdots, q_{k+2}$ are on an S_{k-1}, this S_{k-1} together with the S_{k-1} on L and q_{k+3}, q_{k+4} constitute a quadric which must contain q_1, whence in the projection p_1, p_{k+3}, p_{k+4} are on a line. Here the isolated position of p_1 is not material.

The above discussion suggests the following construction for the set in S_k when the set in the plane is given.

THEOREM 4. *Given the set P_{k+4}^2, let the parameter t of the line pencil on p_1 be*

introduced as a parameter on the linear system Σ_1 *of* ∞^{k-3} *rational curves of order* k *with a* $(k-1)$*-fold point at* p_1. *Then* $t_2, \cdots, t_{k+4}$ *are the parameters of* $p_2, \cdots, p_{k+4}$ *on every curve of* Σ_1 *and the parameters of the multiple point* p_1 *determine a linear system of* ∞^{k-3} *binary* $(k-1)$*-ics all of which are apolar to a binary* k*-ic,* γ_1^k. *In* S_k *select a parameter system* t *on a norm curve* N_1^k. *Then the points of* N_1^k *with parameters* $t_2, \cdots, t_{k+4}$ *and the point of* S_k *determined by* γ_1^k *with reference to* N_1^k *constitute a set* $q_2, \cdots, q_{k+4}$, q_1 *associated with* P_{k+4}^2.

This is indeed an immediate consequence of the fact that the curves of Σ_1 are the projections of N_1^k from the ∞^{k-3} spaces L. This same projection and the further fact that the choice of a single curve of the system Σ_1 is sufficient to determine the corresponding L lead to the following theorem, which is not readily apparent from the plane figure alone.

THEOREM 5. *The* $k+4$ *systems* Σ_i *of dimension* $k-3$ *of rational curves of order* k *with a* $(k-1)$*-fold point at* p_i *and simple points at the remaining points of* P_{k+4}^2 *are in one-to-one correspondence with each other.*

We shall see in § 4 that for Q_7^3 the $\infty^0 = 1$ space L is the point common to all of the ∞^2 elliptic quartics on Q_7^3; for Q_8^4 the ∞^1 spaces L are the common bisecants of all the ∞^1 elliptic quintics on Q_8^4; and for Q_9^5 the ∞^2 spaces L are the trisecant planes of the unique elliptic sextic on Q_9^5. For further sets no equally simple characterization of the spaces L has been obtained.

3. Mapping of P_n^3 upon its associated Q_n^{n-5}

In order to map a set P_8^3 upon its associated Q_8^3 we need only to find a further set P_6^3 such that the set $P_{14}^3 = P_8^3 + P_6^3$ shall have the property that the linear system Σ of cubic surfaces on the 14 points shall have the dimension 6, i.e., that all the cubic surfaces on 13 of the points shall pass through the 14th. For then if 4 of the points of P_8^3 are in a plane π a cubic surface of the system Σ can be made to pass through 6 more points of π in general position and therefore to contain π as a factor. The remaining factor is a quadric on P_6^3 which contains the other four points of P_8^3. Hence the linear system of quadrics on P_6^3 will map P_8^3 upon its associated Q_8^3.

One symmetrical set of 14 points of such character may be obtained as follows. Given 6 points $r_1, \cdots, r_6$ of a plane, select a quartic curve with simple points at r and an octavic curve with triple points at r. These two curves meet elsewhere in 14 points. They are mapped from the plane by cubic curves on the points r into two space sextics of genus 3 with 14 common points. The two space sextics are on one cubic surface—the map of the plane —and only one since the two sextic curves could not lie at once on two cubic surfaces one of which is non-degenerate. Since each sextic curve is on 4 linearly independent cubic surfaces, there must be on their 14 common points $4 + 4 - 1 = 7$ linearly independent cubic surfaces and the set has the required property.

The trisecant locus of the one sextic—an octavic surface with the sextic as a triple curve—meets the other sextic in $8 \times 6 - 14 \times 3 = 6$ points, whence six trisecants of each curve are secants of the other and these two sets of trisecants are a double six of the unique cubic surface on both sextics—the double six of the mapping system. The rôles of the two plane curves are interchanged by the plane Cremona transformation of order 5 with double F-points at the six points r. We observe that the pair of space sextics is the complete intersection of a cubic and a quartic surface.

The number of absolute constants is 4 for the points r and 8 more for each of the plane curves, or 20 in all. Hence in space such a set of 14 points has $20 + 15 = 35$ projective constants. A space sextic of genus three has $15 + 9 = 24$ projective constants so that on a given sextic there are ∞^{11} such sets of 14 points which lie in a linear series I_{11}^{14}. From this there follows that at most 11 of the 14 points can be chosen at random in space. For such sets from P_8^3 to P_{11}^3 we have

THEOREM 6. *The three-dimensional sets P_8^3, P_9^3, P_{10}^3, and P_{11}^3 can be mapped upon their associated sets Q_8^3, Q_9^4, Q_{10}^5, and Q_{11}^6 by the linear system of quadrics on a supplemental set P_6^3, P_5^3, P_4^3, and P_3^3 respectively, which with the given set makes up the 14 points of intersection of two space sextics of genus three.*

The mapping system of this theorem is more general than is needful for the purpose. Consider for example the set P_8^3. It lies on a unique elliptic quartic E^4, the intersection of quadrics Q_1, Q_2. Let C be a cubic surface on P_8^3 which cuts E^4 in a residual set P_4. Let two other points in general position be a set P_2. The totality of cubic surfaces on the 12 points $P_8^3 + P_4$ is made up of $C + \pi Q_1 + \pi' Q_2$ where π, π' are arbitrary planes. In this system of ∞^8 surfaces there is a system of dimension 6 on $P_8^3 + P_4 + P_2$, whence quadrics on $P_4 + P_2$ map P_8^3 upon its associated set Q_8^3. This mapping is however a degenerate case of Theorem 6, since E^4 and a bisecant of E^4 from each point of P_2 make up a degenerate sextic of genus three.

The simplest transition from P_8^3 to Q_8^3 is obtained by taking P_8^3 on an E^4 with canonical parameter u (i.e., such that the coplanar condition is $u_1 + u_2 + u_3 + u_4 \equiv 0 \bmod. \omega_1, \omega_2$) for which the parameters of the points of P_8^3 are $u_1, \cdots, u_8$, where $\Sigma_1^8 u = \sigma$. If now we set $u_i + v_i = \sigma/4$ $(i = 1, \cdots, 8)$ then $v_1 + \cdots + v_4 \equiv \sigma - (u_1 + \cdots + u_4) \equiv u_5 + \cdots + u_8$. Hence the four points v are on a plane if the complementary four points u are on a plane, or the set v is associated to the set u. The lines joining u_i, v_i are generators of a regulus on E^4. For given P_8^3 the $\sigma/4$ has 16 determinations, whence

THEOREM 7. *For a given set P_8^3 there are 16 reguli on the E^4 through P_8^3 such that the generators of a regulus on the points of P_8^3 meet the E^4 again in the points of an associated Q_8^3.*

Again let the set P_9^3 be on a quadric with generators t, τ and let $(a\tau)^2 (\alpha t)^3$

$= 0$ and $(b\tau)^3(\beta t)^2 = 0$ be two quintics of genus two of different kinds on P_9^3 and Q. These quintics meet in four other points P_4 on Q. Let P_1 be a point in general position. Then if C_1, C_1' are cubic surfaces on the first quintic, C_2, C_2' cubic surfaces on the second quintic, and π is an arbitrary plane we have in $\lambda_1 C_1 + \lambda_2 C_1' + \lambda_3 C_2 + \lambda_4 C_2' + \pi Q$ a system of ∞^7 cubic surfaces on P_9^3 and P_4. Hence there will be a system of dimension 6 on P_9^3, P_4, P_1, or the system of quadrics on $P_4 + P_1$ will map P_9^3 upon Q_9^4. This again is a special case of Theorem 6 since a bisecant to the one quintic from P_1 makes up with the quintic a degenerate sextic of genus three and the two sextics thus made up have 14 common points. We shall however find in § 4 a different mode of transition from P_9^3 to Q_9^4 which exhibits more effectively their mutual relations.

There appears to be no point in S_4 from which a general set Q_9^4 can be projected into its associated set. If Q_9^4 is on an elliptic quintic E^5 (two conditions) a quadric on Q_9^4 will cut E^5 in a tenth point from which the desired projection can be made (§ 4, Theorem 11). However, no general sets except planar sets have been found which are the projections of their associated sets. On the other hand no proof of the impossibility of such a projection has been found.

We complete the mapping of sets P_n^3 upon their associated sets by means of an apparatus derived from the elliptic curves. Let E_k^m be an elliptic curve of order $m > k$ in an S_k. It is the projection of the normal E_{m-1}^m from an S_{m-k-2}. The E_{m-1}^m has one absolute constant and the S_{m-k-2} in S_{m-1} has $(m-k-1)(k+1)$ further constants, so that the projection has $(m-k-1)(k+1)+1$ absolute constants. This number added to the $(k+1)^2 - 1$ constants of a projectivity in S_k furnishes $m(k+1)$. Hence the elliptic m-ic in S_k, E_k^m, has $m(k+1)$ constants and can be passed through $[m(k+1)/(k-1)]$ points in S_k, where the bracket indicates the largest integer equal to or less than the number within it.

Since r-ic spreads cut the E_k^m in an I_{mr-1}^m, an r-ic spread on mr general points of E_k^m contains it completely. Hence there are $\infty^{\binom{r+k}{k}-mr-1}$ r-ic spreads on E_k^m and there are $\infty^{\binom{r+k}{k}-mr}$ r-ic spreads on the mr points cut out on E_k^m by a definite r-ic spread.

Beginning then with a set P_{2j}^3 we can pass an E_3' through its points. Let an r-ic surface on P_{2j}^3 meet E_3' in $j(r-2)$ further points $P_{j(r-2)}$. Then there are $\infty^{\binom{r+3}{3}-jr}$ r-ic surfaces on $P_{2j}^3 + P_{j(r-2)}$. If we suppose that these surfaces are subject to $\alpha \geqq 0$ further linear conditions, say to pass through a set of points P_α, we have a linear system of $\infty^{\binom{r+3}{3}-jr-\alpha}$ r-ic surfaces on the base $P_{2j}^3 + P_{j(r-2)} + P_\alpha$. If 4 points of P_{2j}^3 are on a plane and if $\binom{r+3}{3} - jr - \alpha = \binom{r+2}{2} - 4$, then an r-ic surface of the linear system can be determined which contains this plane as a factor leaving an $(r-1)$-ic surface on $P_{j(r-2)} + P_\alpha$

which passes through the remaining $2j-4$ points of P^3_{2j}. This condition becomes

$$\binom{r+2}{3} - jr + 4 = \alpha. \tag{2}$$

Since $\alpha \geqq 0$, then, for given j, r is defined by the inequality

$$\binom{r+2}{3} + 4 \geqq jr. \tag{3}$$

The modification for an odd set P^3_{2i-1} is readily made and we state at once

THEOREM 8. *Through a given set $P^3_{2j}\{P^3_{2j-1}\}$ pass an E^r_3 and cut it by an r-ic surface on $P^3_{2j}\{P^3_{2j-1}\}$ which meets E^r_3 again in a set $P_{j(r-2)}\{P_{j(r-2)+1}\}$ where r is the smallest integer defined by* (3). *The linear system of surfaces of order $r-1$ on this residual set and on a further general set P_α, where α is defined by* (2), *maps S_3 upon a 3-way in $S_{2j-5}\{S_{2j-6}\}$ in such a way that the set $P^3_{2j}\{P^3_{2j-1}\}$ is mapped upon its associated $Q^{2j-5}_{2j}\{Q^{2j-6}_{2j-1}\}$.*

For the sets P^3_9 and P^3_{10} the numbers j, r, α are 5, 4, 4; for P^3_{11} and P^3_{12}, 6, 4, 0; for P^3_{13} and P^3_{14}, 7, 5, 4; etc.

4. Particular and special sets of points

It is the aim in the present section to consider in more detail the relation of particular sets P^k_n for values of n from 8 to 10 to their associated sets both for cases when the n points of the set are in general position and for cases when they are subject to certain conditions. A question naturally arises as to what types of conditions would be most interesting and as to what types of configurations connected with the associated sets would best exhibit the relations sought. In answer to this inquiry we recall the noteworthy theorem in regard to associated sets (P. S., II (16), p. 361), which states that if P^k_n and P'^k_n are congruent under regular Cremona transformation in S_k their associated sets Q^{n-k-2}_n and Q'^{n-k-2}_n are also congruent under regular Cremona transformation in S_{n-k-2}. More specifically, if P^k_n is congruent to P'^k_n under the Cremona involution $x'_i = 1/x_i$ $(i = 1, \cdots, k+1)$ with its $k+1$ F-points at points of P^k_n, then Q^{n-k-2}_n is congruent to Q'^{n-k-2}_n under the involution $x'_i = 1/x_i$ $(i = 1, \cdots, n-k-1)$ with its $n-k-1$ F-points at the complementary $n-k-1$ points of Q^{n-k-2}_n. The regular Cremona group is generated by this one Cremona involution and projectivities.

We shall seek therefore to express the desired relations in terms of such loci or in terms of such properties of these loci as are invariant under regular Cremona transformation. Thus a rational curve, or an elliptic curve, of order $k+1$ on the points of P^k_n is transformed by regular transformation into a curve of the same order on the points of the congruent set. The same is true of multiples of such curves, i.e., curves of orders $l(k+1)$ with l-fold points at the points of P^k_n, if such curves exist. This property of invariance is shared by a certain type of surface—the rational M'_2 in S_{r+1}. We shall first derive some facts concerning this surface for later use.

If $r = 2l + 1$ [$2l$] the system of rational plane curves of order $l + 1$ on the base O^l [O^l, σ] has the dimension $r + 1$ and maps the plane upon a 2-way of order r, M_2, in S_{r+1}. Each of these surfaces is the projection of the one of next higher order from one of its points. This is evidently the case in passing from the base O^l to the base O^l, σ. But also the base O^l, σ, σ' can be reduced by quadratic transformation to the base O^{l-1}. Thus the series of surfaces M_2 constitute the progenitors of the quadric M_2^3 in S_3. Lines on the point O map into the ∞^1 "generators" of the surface.

In case r is odd directions at O map into a unique "directrix," a rational norm curve of order l; while the lines of the plane map into ∞^2 "directors," rational norm curves of order $l + 1$. Since S_r's on the directrix are mapped by sets of $l + 1$ lines on O, and S_r's on a given director by sets of l lines on O and a given line, the directrix and a director are in skew S_l, S_{l+1}, and the generators are lines joining corresponding points of these two rational curves. Included, however, among the ∞^2 directors are the ∞^1 which consist of the fixed directrix and a variable generator.

In case r is even there are ∞^1 directrices, the maps of lines on σ, which are rational norm curves of order l. Included in this system is one curve which is the map of directions at O. As before the ∞^1 generators are the maps of lines on O but this system includes the one line which is the map of directions at σ. The line $\overline{O\sigma}$ is mapped into directions on the surface about the point where the generator σ meets the directrix O. If π, ρ are two lines on σ the mapping system can be expressed in the form $\pi\Sigma_1 + \rho\Sigma_2$ where Σ_1, Σ_2 each is the system of l lines on O. Hence any two of the directrices lie in skew S_l's and the generators are lines joining corresponding points on the two.

In either case by estimating the number of constants involved in the choice of the skew spaces; in the choice of the rational curve in each; in the projectivity between the two curves set up by the generators; and by allowing for the freedom in the choice of the skew spaces for given surface, we find that the number of projective constants of the M_2 is $(r + 2)^2 - 7$, whence the M_2 admits a 6-parameter collineation group. This group for r odd is the map of the 6-parameter collineation group of the plane with fixed point O; for r even it is the map of the 6-parameter quadratic group with fixed F-points at O, σ.

Since it is $r - 1$ conditions that an M_2 in S_{r+1} be on a point, we see that there are ∞^2 M_2's on $r + 5$ points in general position. Thus on 8 points in S_4 there are ∞^2 M_2^3's, or on 9 points a finite number; on 9 points in S_5 there are ∞^2 M_2^4's which fill up a spread, whence for 10 points in S_5 there is a single condition invariant under regular Cremona transformation which expresses that the 10 points lie on an M_2^4.

The system of plane rational curves of order l on the base O^{l-1} [O^{l-1}, σ]

has the dimension $r-1$. Let C_i $(i=1, \cdots, r)$ be linearly independent in this system and let π, ρ be two lines on O. If then we set $m_i = \pi C_i$, $n_i = \rho C_i$, where m_i, n_i are the linear forms in S_{r+1} which cut M_2' in the maps of the given plane curves, we find that the equation of M_2' is

$$\begin{Vmatrix} m_1 & m_2 & \cdots & m_r \\ n_1 & n_2 & \cdots & n_r \end{Vmatrix} = 0. \tag{4}$$

Conversely a manifold in S_{r+1} defined by such a matrix is in general an M_2' mapped as above.

In the case $r = 2l$ a parametric equation of M_2' is

$$x_0 = (\alpha_0 t)(a_0 \tau)^l, \quad x_1 = (\alpha_1 t)(a_1 \tau)^l, \quad \cdots, \quad x_{r+1} = (\alpha_{r+1} t)(a_{r+1} \tau)^l. \tag{5}$$

For given τ we have one of the ∞^1 generators; for given t one of the ∞^1 directrices. In the case $r = 2l+1$ the parametric equation is

$$x_0 = (\alpha_0 t)(a_0 \tau)^{l+1}, \qquad x_1 = (\alpha_1 t)(a_1 \tau)^{l+1}, \qquad \cdots, \qquad x_{r+1} = (\alpha_{r+1} t)(a_{r+1} \tau)^{l+1}, \tag{6}$$

where

$$(\alpha_i t')(a_i \tau')^{l+1} = 0 \qquad (i = 0, \cdots, r+1).$$

This is in fact the projection of (5) for $r = 2l+2$ from a point t', τ' upon it.

Special cases of these rational surfaces occur. Thus cubic curves on the base O^2, o, o' map the plane upon an M_2^3 in S_4. This mapping system can be reduced to conics on the base O by quadratic transformation with F-points at O, o, o' unless o, o' coincide with O in two distinct directions. Thus cubics with node at O and fixed nodal tangents determine an M_2^3 in S_4 which is more properly the projection of an M_2^5 in S_6 from two points on its directrix conic. This special M_2^3 is obtained in S_4 by joining a point directrix to a cubic curve director. Unless expressly mentioned special M_2's of such types will not be considered.

We shall now prove

THEOREM 9. *An M_2' in S_{r+1} is transformed by the Cremona involution $x_i' = 1/x_i$ $(i = 1, \cdots, r+2)$ with $r+2$ F-points on the M_2' into an M_2''.*

The space x' is mapped in the involution upon the space x by the system of spreads of order $r+1$ with r-fold points at the F-points which are the maps from the plane of the points $p_1, \cdots, p_{r+2}$. Then, for $r = 2l+1$, the transform of M_2' is mapped from the plane by curves of order $2(l+1)^2$ with a $2l(l+1)$-fold point at O and $(2l+1)$-fold points at $p_1, \cdots, p_{2l+3}$; for $r = 2l$, by curves of order $(l+1)(2l+1)$ with an $l(2l+1)$-fold point at O, a $(2l+1)$-fold point at σ; and $2l$-fold points at $p_1, \cdots, p_{2l+2}$. We have merely to show that the two latter mapping systems can be transformed by ternary Cremona transformation into systems of order $l+1$ on the bases O^l or O^l, σ respectively. For odd r this transformation is effected by using first the Jonquière transformation J^{l+1} of order $l+1$ with l-fold point (*center*)

at O and simple F-points at $p_1, \cdots, p_{2l}$, then a quadratic transformation with F-points at p_{2l+1}, p_{2l+2}, p_{2l+3}, and finally the transformation J^{l+1} again. For even r we use first a quadratic transformation with F-points at O, σ, p_1, then a J^l with center at O and simple F-points at $p_2, \cdots, p_{2l-1}$, then the quadratic transformation with F-points at p_{2l}, p_{2l+1}, p_{2l+2}, and finally J^l again. It is easily verified that these transformations effect the required change in the mapping system and the proof is complete.

The three theorems which follow relate to special sets of points when, for given k, n is sufficiently large.

THEOREM 10. *If P_n^k is on a rational norm curve N^k in S_k, then its associated Q_n^{n-k-2} is on a rational norm curve N^{n-k-2} in S_{n-k-2}. The n parameters of the two sets on their respective norm curves are projective. If $(n-k-2)-k = l+1>0$, the set Q is projected upon the set P from any one of the ∞^{l+1} spaces L_l which are $(l+1)$-secant to N^{n-k-2}.*

THEOREM 11. *If P_n^k is on an elliptic norm curve E^{k+1} in S_k, then its associated Q_n^{n-k-2} is on an elliptic norm curve E^{n-k-1} in S_{n-k-2}. If $(n-k-2)-k = l+1>0$, the set Q is projected upon the set P from any one of the ∞^l spaces L_l which are $(l+1)$-secant to E^{n-k-1} at the $l+1$ points cut out by a quadric on Q.*

THEOREM 12. *If P_n^k is on a rational norm surface M_2^{k-1} in S_k, then its associated Q_n^{n-k-2} is on a rational norm surface N_2^{n-k-3} in S_{n-k-2}. Then parameters of the two sets of generators on the points are projective.*

In Theorem 10 let the norm curves in S_k and S_{n-k-2} have the respective parametric equations

$$x_0 = 1, \quad x_1 = t, \quad \cdots, \quad x_k = t^k;$$

$$x_0 = 1, \quad x_1 = t, \quad \cdots, \quad x_{n-k-2} = t^{n-k-2};$$

and let the sets P_n, Q_n be determined on these curves by the parameters $t_1, \cdots, t_n$. If $\lambda_1, \cdots, \lambda_n$ are determined by the $n-1$ equations

$$\lambda_1 t_1^i + \lambda_2 t_2^i + \cdots + \lambda_n t_n^i = 0 \qquad (i = 0, 1, \cdots, n-2),$$

then the points of the one set, affected respectively by factors of proportionality $\lambda_1, \cdots, \lambda_n$, satisfy with the points of the other the bilinear relations requisite for association. We observe that here P_n^2 is obtained by projection of Q_n^{n-4} from ∞^{n-6} spaces L_l rather than ∞^{n-7} spaces as in the general case of Theorem 3.

In Theorem 11 let the canonical parameters of P_n^k on E^{k+1} be $u_1, \cdots, u_n$ where $u_1 + \cdots + u_n + b \equiv 0$. Choose then a mapping system on a base B such that the members meet E^{k+1} in $n-k-1$ variable points and also in a certain number of fixed points whose parameters sum up to b. Then, if $k+1$ points of P_n^k are on an S_{k-1}, $u_1 + \cdots + u_{k+1} \equiv 0$ and $u_{k+2} + \cdots + u_n + b \equiv 0$, whence the complementary $n-k-1$ points of P_n^k are on a member of the mapping system or the $n-k-1$ points of Q_n^{n-k-2}, mapped

from P_n^k, are on an S_{n-k-3}. Thus E^{k+1} is mapped upon E^{n-k-1} and P_n^k is mapped upon its associated Q_n^{n-k-2}. In this way we find upon each of the associated sets P_7^2, Q_7^3, ∞^2 elliptic norm curves, upon each of the associated sets P_8^2, Q_8^4, ∞^1 elliptic norm curves, and upon each of the associated sets P_9^2, Q_9^5, a unique elliptic norm curve.

For Theorem 12 we give the details of the proof only for the case $k = 2l + 1$. Then M_2^{k-1} is the map of the plane by curves of order $l + 1$ on the base O^l, σ and P_n^k is the map of a set π_n^2 in the plane. If $n = 2m + 1$ then O is the center and σ, π_n^2 the simple F-points of a J^{m+2} whose inverse center and F-points are O', σ', $\pi_n'^2$. Curves of order $m - l - 1$ on the base O'^{m-l-2} map the plane on an M_2^{n-k-3} and map $\pi_n'^2$ upon a set Q_n^{n-k-2} which is associated to P_n^k. For if $k + 1$ points $p_1, \cdots, p_{2l+2}$ of P_n^k are on an S_{k-1} there is a curve of order $l + 1$ with l-fold point at O and simple points at $\sigma, \pi_1, \cdots, \pi_{2l+2}$. This curve is transformed by J^{m+2} into a curve of order $m - l - 1$ with $(m - l - 2)$-fold point at O' and simple points at $\pi'_{2l+3}, \cdots, \pi'_{2m+1}$. Hence the points $q_{2l+3}, \cdots, q_n$ are on an S_{n-k-3} in S_{n-k-2}. If, however, n is even we take σ, σ' to be a pair of ordinary corresponding points for a $J^{1+\frac{n}{2}}$.

It should be observed however that an M_2^2 in S_3, an ordinary quadric, counts in two ways as a ruled normal surface. It is mapped from the plane by conics on O, σ and as the points are interchanged in the above proof two normal surfaces in S_{n-k-2} are obtained. Hence

THEOREM 13. *If P_n^3 is on a quadric surface which is not a cone, its associated Q_n^{n-5} is on two normal M_2^{n-6}s in S_{n-5}.*

A simple statement of the relations among the F-points of a Jonquière transformation can be given in terms of associated sets.

THEOREM 14. *Given the Jonquière transformation J^{n+1} with center at p and simple F-points at P_{2n}^2, then curves of order $n - 2$ with an $(n - 3)$-fold point at p map the plane upon an M_2^{2n-5} in S_{2n-4} and map the set P_{2n}^2 upon a set R_{2n}^{2n-4} which is associated to the set Q_{2n}^2 of simple F-points of the inverse transformation.*

The proof of this is immediate by the foregoing methods.

We now proceed to particular sets beginning with P_8^2, Q_8^4. The ∞^1 elliptic quintics, E^5's, on Q_8^4 are obtained by the mapping of P_8^2 on Q_8^4 by conics on the 9th base point p_9 of the pencil of cubics on P_8^2. This pencil becomes a pencil of E^5's on an M_2^3 on Q_8^4 and the generators of M_2^3, which arise from the lines of the plane on p_9, are bisecants of all these E^5's. However, each of the ∞^1 E^5's on Q_8^4 has ∞^1 M_2^3's on it, whose generators on points r_1, r_2 satisfy the involution $r_1 + r_2 \equiv k$.* That particular M_2^3 common to all the E^5's is determined by the involution cut out on any E^5 by quadrics on Q_8^4. For if, in the plane, $u_1 + u_2 + \cdots + u_8 + u_9 \equiv 0$, $v_1 + \cdots + v_5 + u_9 \equiv 0$, $w_1 + w_2 + u_9 \equiv 0$ represent the sections of a cubic of the pencil by respectively a cubic

*Segre, Mathematische Annalen, vol. 27 (1886).

of the pencil, a mapping conic, and a line on p_9, then, on writing the second relation in the form $(v_1 + \frac{1}{5}u_9) + \cdots + (v_5 + \frac{1}{5}u_9) \equiv 0$ in order to introduce the canonical parameter $v' = v + \frac{1}{5}u_9$ on the mapped E^5, we have for Q_8^4 and the meets of a generator of the unique M_2^3 the relations

$$(u_1 + \tfrac{1}{5}u_9) + \cdots + (u_8 + \tfrac{1}{5}u_9) \equiv \tfrac{8}{5}u_9, \quad (w_1 + \tfrac{1}{5}u_9) + (w_2 + \tfrac{1}{5}u_9) \equiv -\tfrac{3}{5}u_9,$$

whence on E^5 $v_1' + \cdots + v_8' + w_1' + w_2' \equiv 0$ and the ten points are a quadric section.

We may relate Q_8^4 and any one of the ∞^2 M_2^3's on it to P_8^2 in the plane as follows. Let P_8^2, R_8^2 be F-points of a J^5 with centers at p, r where p is any one of the ∞^2 points of the plane. Then if p_1, p_2, p_3 are on a line, the points r_4, $\cdots$, r_8, r are on a conic. Hence conics on r map the plane upon an M_2^3 in S_4 in such a way that R_8^2 is mapped upon the set Q_8^4 associated to P_8^2.

In addition to the ∞^1 E^5's on Q_8^4 there are ∞^2 rational quintics R^5 on Q_8^4. These are in one-to-one correspondence with the ∞^2 M_2^3's on Q_8^4. For, given an M_2^3 on Q_8^4, of the 7 linearly independent quadrics on Q_8^4 three are on M_2^3 (the three determinants of the matrix (4)) and of the remaining four one is on the directrix of M_2^3 and cuts M_2^3 in a residual R^5 trisecant to the directrix and unisecant to the generators. Conversely, given an R^5 on Q_8^4 it has a unique trisecant (with parameters determined by the canonizant of the binary quintic apolar to all S_3 sections) whose points are in 1–1 correspondence with the points of the curve (the correspondence being determined by making the three points common to the curve and trisecant self-corresponding) and the lines joining corresponding points are generators of an M_2^3 on Q_8^4. The question then arises as to the nature of the spread which is the locus of the ∞^2 R^5's on Q_8^4 or the nature of the condition that a Q_8^4 be on an R^5, and as to the corresponding condition on the associated P_9^3. The two theorems which follow answer these questions.

THEOREM 15. *There are two M_2^3's on a given Q_9^4 which are covariantly related to the set under regular Cremona transformation. They are isolated by the same irrationality as separates the two reguli on the unique quadric on the associated set P_9^3. The parameters of the 9 generators of one of the M_2^3's on Q_9^4 are projective to those of the 9 generators of one of the reguli on P_9^3. If the set Q_9^4 lies on an R^5 (a single condition) then it lies on but one M_2^3 and its associated P_9^3 lies on a quadric cone.*

Two M_2^3's in S_4 meet in a set Q_9^4. That on Q_9^4 there are two M_2^3's is proved by Theorem 13. That there are only two is proved as follows. The ∞^2 M_2^3's on Q_8^4 are loci of ∞^1 bisecants of the ∞^1 E^5's on Q_8^4. One of these M_2^3's, say m_2^3, is a locus of bisecants of each of the E^5's; the others are each a bisecant locus of only one E^5. If then an M_2^3 is on a 9th point q_9 there is a bisecant of an E^5 on q_9; if two M_2^3's are on q_9 their plane cuts m_2^3 in the 4 meets of two

bisecants with their respective E^5's. Hence this plane cuts m_2^3 in one of the director conics on it. A third bisecant on Q_9 would have to be in this plane else there would be two director conics with two intersections, whereas such conics have only one. A director conic meets each of the ∞^1 E^5's in three points and on this conic there is an involution of triads whose joining triangles envelop another conic. Hence on the point q_9 in the plane of this conic there are just two lines of this envelope each belonging to one of the two M_2^3's on Q_9^4.

If Q_9^4 is on an R^5 which must lie on one of the two M_2^{3}'s on Q_8^4 and must cut its directrix in three points and each generator in one point, then in the notation of the proof of Theorem 12 the R^5 must be the map of a rational plane quartic with triple point at O' and on π_8', $\cdots$, π_8' as well as on σ'. But then σ must coincide in some direction with O, and the quadric on P_9^3 mapped by conics on O, σ is a quadric cone.

An E^5 in S_4 is projected from a line into an elliptic plane quintic with five nodes and from a line which meets E^5 into an elliptic plane quartic with two nodes, whence the bisecant locus of E^5 is a quintic spread on which E^5 is a triple curve. The ∞^1 E^5's on a given Q_8^4 can be put into 1, 1 correspondence with a pencil of plane cubics and therefore can be named rationally in terms of a parameter λ. Through a point there pass two bisecants belonging to two of these E^5's, whence the aggregate of these bisecant spreads of the ∞^1 E^5's constitute a quadratic system. The two bisecants isolate the two M_2^3's on Q_8^4 and the given point, whence if they coincide the two M_2^3's coincide and the given point and Q_8^4 are on an R^5. Hence

THEOREM 16. *If* $\lambda^2 B_0 + 2\lambda B_1 + B_2 = 0$ *is the quadratic system of bisecant spreads of the* ∞^1 E^5*'s on* Q_8^4, *the spread* $B_1^2 - B_0 B_2 = 0$ (*a spread of order* 10 *with* 6*-fold points at* Q_8^4 *and a double* M_2^3 *consisting of the* ∞^1 E^5*'s*) *is the locus of the* ∞^2 *rational quintics on* Q_8^4, *or the locus of points through which there can be drawn but one line bisecant to an* E^5 *on* Q_8^4, *or through which there can be passed but one* M_2^3 *on* Q_8^4. *Its equation may be obtained by replacing in the condition that a quadric on* P_9^3 *be nodal* (*a condition of degree* 8 *in the coördinates of each point of* P_9^3 *whose terms consist of products of* 18 *determinants* $|p_r p_j\ p_k p_l|$) *each determinant* $|p_{i_1} p_{i_2} p_{i_3} p_{i_4}|$ *by the complementary determinant* $|q_{i_5} q_{i_6} q_{i_7} q_{i_8} q_{i_9}|$ *formed for* Q_9^4 *and allowing the* 9*th point to vary.*

Here then we have an instance of the actual determination of a covariant of Q_8^4 or an invariant of Q_9^4 under the infinite group of regular Cremona transformations attached to the set.

We complete the discussion of sets of 9 points with the Q_9^5 associated with the set P_9^2. In S_5 the elliptic norm sextic E^6 has one absolute and 36 projective constants; the rational sextic R^6 has three absolute and 38 projective constants; and the M_2^4 has 29 projective constants; whence on Q_9^5 there is a finite number of E^6's, ∞^2 R^6's, and ∞^2 M_2^4's. There is, however, in S_5 a new

type of rational 2-way of order 4, the Veronese surface V_2^4, which shares with M_2^4 the property that its projection from one of its points is an M_2^3. The V_2^4 is the map of the plane by the linear system of all conics in the plane. It contains ∞^2 conics, the maps of lines of the plane, and the locus of the ∞^2 planes of these conics is a V_4^3 upon which V_2^4 is a double manifold. Analytically V_4^3 is obtained by setting a 3-row symmetric determinant of linear forms equal to zero and V_2^4 is the locus for which the six first minors vanish. The V_2^4 is unaltered by an 8-parameter collineation group, the map of the ternary group, whence it has $35 - 8 = 27$ projective constants. We should expect, therefore, to find on Q_0^5 a finite number of V_2^4's. The surface V_2^4 shares with M_2^4 also the property expressed by

THEOREM 17. *The Veronese surface V_2^4 is transformed into a Veronese surface $V_2'^4$ by a regular Cremona transformation whose F-points are on V_2^4. If the regular transformation in S_5 is $y_i = 1/x_i$ $(i = 0, \cdots, 5)$ the two V_2^4's are mapped by conics from planes which are in correspondence under the ternary quintic transformation with 6 double F-points. The V_4^3 with double V_2^4 is transformed into the $V_4'^3$ with double $V_2'^4$.*

Indeed the given involution maps the $S_4(y)$'s upon a system of quintic spreads with 4-fold points at the 6 F-points on V_2^4. This is the map of a ternary system of 10-ics with 4-fold points at 6 points, which can be transformed by the ternary transformation mentioned into a system of conics. The same involution transforms a cubic spread with nodes at the 6 F-points into a similar spread, whence V_4^3 on V_2^4 passes into $V_4'^3$ on $V_2'^4$.

Upon V_2^4 there is a linear system of ∞^9 E^6's, the maps of cubic curves in the plane. Conversely an E^6 on V_2^4 is cut out by a quadric which meets V_2^4 in a residual conic, whence the corresponding quartic in the plane breaks up into a line and a cubic. Therefore there are no other E^6's on V_2^4. The conics on V_2^4 are trisecant to the E^6's on V_2^4. A canonical elliptic parameter on the plane cubic is mapped into a canonical parameter on E^6 whence the planes of V_4^3 are those which meet E^6 in three points for which $u_1 + u_2 + u_3 \equiv 0$. Obviously any two of these planes lie in an S_4 and meet in a point. But the same thing is true of the three other involutions for which $u_1 + u_2 + u_3 \equiv \omega/2$. Hence on E^6 there are 4 V_2^4's or also there are 4 V_4^3's which contain E^6 doubled. Such a V_4^3 must contain every bisecant of E^6. The locus of bisecants, B_3^9, of E^6 is a 3-way of order 9 which has E^6 as a 4-fold curve, since from a plane E^6 is projected into a plane sextic with 9 nodes, and from a plane which meets E^6 the E^6 is projected into a plane quintic with 5 nodes. Hence the bisecant locus is the complete intersection of two of the four V_4^3's and the four lie in a pencil. A member of this pencil other than a V_4^3 also contains B_3^9. Given then a trisecant plane for which $u_1 + u_2 + u_3 \equiv k$, the above pencil of W_4^3's contains the three bisecants in the plane, whence one member, say W_4^3, contains

the plane. Since any plane for which $v_1 + v_2 + v_3 \equiv -k$ meets the above plane in a point, W_4^3 must meet this latter plane in its bisecants and an outside point and therefore must contain it. Hence W_4^3 is the locus of the ∞^2 trisecant planes $v_1 + v_2 + v_3 \equiv -k$ or also of the ∞^2 trisecant planes $u_1 + u_2 + u_3 \equiv k$. For each of the 4 V_4^3's in the pencil of W_4^3's the two systems of generating planes coincide into a single system, since $k \equiv -k$ when $k \equiv \omega/2$. Hence

THEOREM 18. *An E^6 is contained on 4 V_2^{4}'s whose V_4^3's are in the pencil of spreads W_4^3 on the bisecant locus B_3^9 of E^6 for which E^6 is a 4-fold curve. A particular W_4^3 of the pencil with double E^6 has the two systems of ∞^2 generating trisecant planes for which $u_1 + u_2 + u_3 \equiv -k$, k which coincide for the 4 V_4^3's. Under regular Cremona transformation with F-points on E^6 the properties of this pencil are invariant.*

That there are on E^6 four V_2^{4}'s may be seen by the use of an elementary theorem. Isolate one of the V_2^4's as the map of a plane. The V_4^3's of the other three V_2^4's cut the isolated one in E^6 doubled, whence in the plane we have the square of a cubic expressed in three ways as a symmetric 3-row determinant whose elements are conics. But we know that a cubic can be expressed in three ways as a symmetric 3-row determinant of linear forms, since it is the hessian of three cubics and the square of a symmetric determinant is symmetric. Moreover we know that the relation of the hessian to the three cubics involves the three half periods.

THEOREM 19. *On a general set Q_9^5 there is a unique E^6 and four V_2^{4}'s.*

We see at once that an E^6 and an E'^6 on Q_9^5 could not have different absolute invariants. For an E^6 on Q_9^5 is projected from a properly chosen trisecant plane into an E^3 on the associated P_9^2, and E'^6 into an E'^3 on P_9^2, whence, since E^3 and E'^3 cannot coincide, the set P_9^2 is the special set of 9 base points of a pencil of E^3's and Q_9^5 is also a special set. If, however, there were an E^6 and an E'^6 on Q_9^5, then on projecting from q_9 we should have in S_4 an E^5 and E'^5 on R_8^4, members of a pencil on an M_2^3 in S_4. Hence in S_5 there are ∞^1 elliptic quintic 2-way cones with vertex at q_9 and on $q_1, \cdots, q_8$, and with no other points common to any two. A quadric on Q_9^5 and four generators of any one of these cones meets the cone in an E^6 on Q_9^5, whence there is a pencil of such E^6's on Q_9^5 with all values of the absolute invariant and again Q_9^5 is the special set above. This unique E^6 and therefore Q_9^5 also carries four V_2^4's. There are no V_2^4's on Q_9^5 which are not also on E^6, else there would be on such a V_2^4 an E'^6 on Q_9^5.

THEOREM 20. *If P_9^2 is the set of base points of a pencil of E^3's, its associated Q_9^5 is the set of base points of a pencil of E^6's on a V_2^4, the map of the plane by conics.*

This is an immediate consequence of the elementary theorem that if three of the points of such a planar set are on a line the remaining six are on a conic.

We observe that for such a pencil of E^6's on a given V_2^4 each E^6 according to Theorem 18 is contained on three other V_2^4's whence this special Q_9^5 is on ∞^1 V_2^4's one of which is isolated while the others divide into triads which depend rationally on a parameter.

If 8 points of such a special Q_9^5 are given, the locus of the 9th is a 3-way, four of whose points are on any E^6 through the given 8 points. This 3-way is the extension of the Weddle surface and bears the same relation to the hyperelliptic functions of genus three as the Weddle to those of genus two. This relation will be discussed in a forthcoming paper.

If P_9^2 of Theorem 20 is the set of flex points of an E^3, the base points of a syzygetic pencil, then any two are on a line with a third, whence

THEOREM 21. *There exists in S_5 a set of 9 points invariant under a Hesse collineation G_{216} with the property that any two points determine a third such that the remaining six are on an S_4. The configuration contains 12 S_4's, eight on each point.*

This set of 9 points has the unusual property that if six be selected which form a reference 6-point, no other one can be taken to be the unit point, since each of the other three must lie in one of the reference S_4's. Using a proper set of six as reference points the coördinates of the other three are

$$\begin{array}{cccccc} \omega, & \omega^2, & -1, & -\omega^2, & 0, & -\omega; \\ -1, & 1, & \omega^2, & -1, & -\omega^2, & 0; \\ 1, & -1, & \omega, & 0, & -\omega, & -1 \end{array} \qquad (\omega = e^{2\pi i/3}).$$

The problem of obtaining the four surfaces V_2^4 on a given Q_9^5 may be solved through the use of the associated set P_9^2 as follows:

THEOREM 22. *On the E^3 on P_9^2 join the 9th base point of the pencil on p_1, $\cdots$, p_8 to p_9 to meet E^3 again in p'. From p' draw a tangent to E^3 at p'' (4 choices). Construct a set $r_9, r_1, \cdots, r_8$ congruent to $p'', p_1, \cdots, p_8$ under J^3 with centers r_9, p''. Then conics map the set R_9^2 upon the set Q_9^5 associated to P_9^2 and map the plane upon one of the four V_2^4s on Q_9^5.*

We now consider sets of 9 and of 10 points in S_5 with reference to the normal surfaces M_2^4 and the rational sextic curves R^6. We have noted that on Q_9^5 there are ∞^2 M_2^4's and ∞^2 R^6's. Only ∞^1 of the M_2^4's contain the unique E^6 on Q_9^5. For the M_2^4 mapped from the plane by cubic curves on the base O^2, σ contains ∞^8 E^6's which are mapped from quartic curves with nodes at O, σ whence M_2^4 and E^6 on it have 37 constants. But E^6 alone has 36 constants, whence on E^6 there are ∞^1 M_2^4's. These are the bisecants of the ∞^1 involutions $u + u' \equiv k$, since lines on O cut out such an involution on a ternary quartic with node at O. An M_2^4 on Q_9^5 and not containing E^6 can have no other point in common with E^6. For if E^6 were to meet M_2^4 in 10 points, at least four of

the quadrics on M_2^4 would contain E^6. But four such quadrics meet in a residual conic. We now prove

THEOREM 23. *The locus of the ∞^2 M_2^4's on Q_9^5 is a cubic spread with the E^6 on Q_9^5 for double curve. A point of this cubic spread forms with Q_9^5 a symmetrical set Q_{10}^5 which are the meets of two M_2^4's and whose associated set P_{10}^3 is on a quadric surface. The cubic spread is that locus of ∞^2 trisecant planes of E^6 whose meets with E^6 lie with Q_9^5 on a quadric.*

For the condition that P_{10}^3 is on a quadric surface is of degree two in each point p_i and therefore is a sum of products of 5 four-row determinants. The corresponding condition on Q_{10}^5 is a sum of products of 5 six-row determinants and therefore is of degree three in each point q_i. Since the condition on P_{10}^3 is invariant under regular Cremona transformation this is likewise true of Q_{10}^5. Hence if q_{10} is variable the cubic spread must have nodes at Q_9^5. According to Theorem 13, Q_{10}^5 is the set of points of intersection of two M_2^4's on Q_9^5. Since the cubic spread contains the ∞^1 M_2^4's on Q_9^5 which contain E^6, it contains the bisecant locus B_3^9 of E^6 and therefore is a member of the pencil of Theorem 18 and contains E^6 as a double curve. To prove the trisecant plane property we observe (and omit the verification) that if a quadric contains M_2^4, a plane on this quadric meets M_2^4 in a point. Given then an M_2^4 and a plane trisecant to E^6 at v_1, v_2, v_3 such that $u_1 + \cdots + u_9 + v_1 + v_2 + v_3 \equiv 0$, of the 6 quadrics on M_2^4 and therefore on Q_9^5 at least four are on v_1, v_2, v_3 and at least one contains the plane $v_1 v_2 v_3$ which therefore meets M_2^4 in a point. As M_2^4 varies in the ∞^2 system on Q_9^5, this point runs over the trisecant plane.

An R^6 on Q_9^5 is on a unique M_2^4 on Q_9^5 and vice versa. For given the M_2^4 mapped by cubics on O^2, σ the R^6's are mapped from ternary quintics with 4-fold point at O and simple point at σ, whence on Q_9^5 there is a unique R^6. These R^6's meet the generators in one point and the directrix conics in four points whose four parameters on conic and on R^6 are projective. Given R^6 on Q_9^5, its quadrisecant planes each carry a unique conic with the projective 4-point property just mentioned and the locus of these conics is the unique M_2^4 on Q_9^5 and R^6.

If in the proof of Theorem 12 the point σ' is on a ternary quintic which maps into an R^6 on Q_9^5, then σ coincides with O in some direction and the set P_{10}^3 is on a nodal quadric. For such a set the two M_2^4's coincide. Hence

THEOREM 24. *The two conditions that Q_{10}^5 be on an R^6 are that its associated P_{10}^3 be on a nodal quadric. On such a Q_{10}^5 there is but one M_2^4.*

The ∞^1 M_2^4's on Q_9^5 which contain the E^6 on Q_9^5 are obtained by mapping P_9^2 on Q_9^5 in the ∞^1 ways described in § 1. All of the ∞^2 M_2^4's on Q_9^5 are obtained by the following construction.

THEOREM 25. *For the set P_9^2 we choose a center p (in ∞^2 ways) and, for arbitrarily chosen p_{10}, construct a set $r_1, \cdots, r_9, \sigma, O$ congruent to $p_1, \cdots,$*

p_9, p_{10}, p *under* J^6 *with centers* O, p. *Then cubics on* $O^2\sigma$ *map* $r_1, \cdots, r_9$ *upon the set* Q_9^5 *associated to* P_9^2, *and map the plane upon one of the* ∞^2 M_2^4*'s on* Q_9^5.

For if p_1, p_2, p_3 are on a line, then $r_4, \cdots, r_9, \sigma, O^2$ are on a cubic or $q_4, \cdots, q_9$ are on an S_4. That these M_2^4's are all distinct follows from the fact that the ∞^2 line pencils from p to P_9^2 are projectively distinct. We observe that, when p has been chosen and thereby an M_2^4 isolated, the variation of p_{10} implies the variation of the point $\overline{\sigma O}$ of M_2^4 over the M_2^4.

We shall close with an application to the sets of 10 nodes of a rational plane sextic and of a symmetroid quartic surface Σ. These two figures are related as follows. The sextic $S(t)$ has a conjugate rational sextic $R(t)$ in space such that the plane sections of the one are apolar to the line sections of the other. The locus of planes which cut $R(t)$ in catalectic sextics is Σ (as an envelope) and the 10 planes which cut $R(t)$ in cyclic sextics (reducible to a sum of two sixth powers) are the ten double planes of Σ. If such a cyclic sextic is $(p_1 t)^6 + (p_2 t)^6 = 0$, then $(p_1 t)\cdot(p_2 t) = 0$ are the nodal parameters of a double point of $S(t)$. Thus the nodes of Σ and the nodes of $S(t)$ are in correspondence. It is known that there are two projectively distinct rational sextics $S(t)$, $S(\tau)$ which determine the same Σ. I have proved but not yet published the fact that if 6 nodes of $S(t)$ are on a conic then the complementary 4 nodes of Σ are on a plane. Hence conics on the plane map the plane on a V_2^4 in S_5 and the ten nodes of $S(t)$ upon a Q_{10}^5 on V_2^4 which is associated to the P_{10}^3 of nodes of Σ. But also conics of the plane of $S(\tau)$ map this plane on a $V_2'^4$ and the ten nodes of $S(\tau)$ upon the same Q_{10}^5, since this set also is associated to P_{10}^2. From this there follows at once

THEOREM 26. *If two Veronese surfaces* V_2^4, $V_2'^4$ *meet in* 10 *points,* Q_{10}^5, *then this set is associated to the set* P_{10}^3 *of nodes of a Cayley symmetroid. The spreads* V_4^3, $V_4'^3$, *with double* V_2^4, $V_2'^4$ *respectively, each cut the double spread of the other in a* 12*-ic curve with nodes at* Q_{10}^5. *These curves are the maps from the plane of the two rational plane sextics associated with the symmetroid.*

UNIVERSITY OF ILLINOIS,
URBANA, ILL.

ON ALGEBRAIC FUNCTIONS WHICH CAN BE EXPRESSED IN TERMS OF RADICALS*

BY

J. F. RITT

I. Introduction

We consider, in this paper, an irreducible algebraic relation

$$F(w, z) = 0,$$

of degree n in w and of genus p, and seek to determine those cases in which w can be expressed in terms of z by means of radicals. Why the results obtained here should not have been found before this time is a question which has puzzled us as much as it will puzzle the reader.

Our results for the case in which n is prime are fairly complete. We show first that *if n is prime, and if w can be expressed in terms of radicals, then for every genus except zero there exists an upper bound for n.* For any given genus the mechanisms of the possible Riemann surfaces for w can be determined in a finite number of steps.

For the case of genus zero our problem becomes that of finding all rational functions whose inverses can be expressed in terms of radicals. Given one such function, others can be obtained by performing linear transformations upon the function and upon the variable. We show that all rational functions of prime degree with inverses expressible in terms of radicals can be thus obtained from functions of the following types:

(*a*) *The powers of w.*

(*b*) *The polynomials which occur in the formulas for the multiplication of the argument in the function* cos u.

(*c*) *The fractional rational functions which occur in the formulas for the transformations of the periods of the function $\wp u$.*

(*d*) *For the case of $n \equiv 1 \pmod 4$, the fractional rational functions met, in the lemniscatic case ($g_3 = 0$), in the formulas for the multiplication of the argument of $\wp^2 u$ by $\alpha \pm \beta i$, where $\alpha^2 + \beta^2 = n^2$.*

(*e*) *For the case of $n \equiv 1 \pmod 6$, the fractional rational functions met, in the equianharmonic case ($g_2 = 0$), in the formulas for the multiplication of the arguments of $\wp' u$ and $\wp'^2 u$ by $\alpha \pm \beta i \sqrt{3}$, where $\alpha^2 + 3\beta^2 = n^2$.*

The functions just described are already known to have inverses expressible

* Presented to the Society, October 30, 1920, and October 29, 1921.

in terms of radicals. We prove here that they are the only such rational functions of prime degree.

What we have done for the case in which n is composite is to determine all polynomials whose inverses can be expressed in terms of radicals. In stating our result we employ the terms of our paper *Prime and composite polynomials.** A polynomial $F(z)$ is there called *composite* if there exist two polynomials, $\phi_1(z)$ and $\phi_2(z)$, each of degree greater than unity, such that $F(z) = \phi_1[\phi_2(z)]$. Otherwise, $F(z)$, if of degree greater than unity, is called prime. Let

$$F = \phi_1 \phi_2 \cdots \phi_r,$$

where each $\phi_i(z)$ is a prime polynomial, and is understood to be substituted for z in the polynomial which precedes it.

We show that *if the inverse of $F(z)$ can be expressed in terms of radicals, each $\phi_i(z)$, if not of degree 4, can be obtained by means of linear transformations either from a prime power of z or from a trigonometric polynomial of prime degree.* We will have, that is, if $\phi_i(z)$ is of degree $m \neq 4$,

$$\phi_i = \lambda_1 \pi \lambda_2,$$

where $\lambda_1(z)$ and $\lambda_2(z)$ are linear and where either $\pi(z) = z^m$ or else $\cos mu = \pi(\cos u)$.

The work for this case, with n composite, consists mainly in the proof of a theorem on substitution groups.

II. Functions with a prime number of values

If the degree n in w of the irreducible algebraic relation

$$F(w, z) = 0 \tag{1}$$

is prime, and if w can be expressed in terms of radicals, the group of monodromy of (1) is either the metacyclic group or one of its transitive subgroups. The group of monodromy must contain a substitution of order n. We number the branches of w in such a way that the group contains the substitution $(1\ 2 \cdots n)$, and represent the metacyclic group with the formula

$$\nu' \equiv a\nu + b \pmod{n} \qquad \begin{pmatrix} a = 1, 2, \cdots, n-1 \\ b = 0, 1, 2, \cdots, n-1 \end{pmatrix}. \tag{2}$$

The non-identical substitutions with $a = 1$ displace every index ν. The substitutions with $a \neq 1$ leave a single index fixed. If a substitution of the metacyclic group consists of more than one cycle, its cycles are all of the same order.

Suppose that w has q critical points. We consider the elementary substitutions of the group of monodromy, which correspond to single turns

* These Transactions, vol. 23 (1922), p. 51.

around these critical points. Suppose that there are α of them with $a = 1$ and $q - \alpha$ with $a \neq 1$. We designate the orders of the latter by $s_1, s_2, \cdots, s_{q-\alpha}$.

If the genus of (1) is p, we have, according to the well known formula of Riemann,

$$(n-1)\alpha + \sum_{i=1}^{i=q-\alpha} \frac{n-1}{s_i}(s_i - 1) = 2(n-1) + 2p,$$

or

$$\sum_{i=1}^{i=q-\alpha} \frac{1}{s_i} = q - 2 - \frac{2p}{n-1}. \tag{3}$$

Since no s_i is less than 2, the first member of (3) is not greater than $q/2$, and we find from (3)

$$q \leqq 4 + \frac{4p}{n-1}, \tag{4}$$

which shows that when p is given, there exists an upper bound for q, independent of n.

We shall prove now that *if p is not zero, there exists an upper bound for n which depends only on p.*

Suppose first that $q > 4$. We have then from (4),

$$n - 1 \leqq \frac{4p}{q-4} \leqq 4p.$$

If $q = 4$, equation (3) gives

$$\frac{2p}{n-1} = 2 - \sum \frac{1}{s_i}.$$

It is seen quickly that the second member of this last equation is at least equal to $1/6$. We have thus, in this case,

$$n - 1 \leqq 12p.$$

If $q = 3$, equation (3) gives

$$\frac{2p}{n-1} = 1 - \sum \frac{1}{s_i}.$$

The three integers the sum of whose reciprocals is less than unity by as small a positive number as possible are 2, 3 and 7. We have thus for this case

$$n - 1 \leqq 84p.^*$$

A closer examination of the problem would lead to smaller bounds for n than those found above.

On being given p, we can, with the help of equation (3) and of the upper

* If $q - \alpha < 3$, we have a stronger inequality.

bounds for n and q, determine the possible Riemann surfaces for w in a finite number of steps. We shall not follow this question further in the present paper.

We consider the case of $p = 0$, which includes the most interesting examples already known of algebraic functions expressible in terms of radicals. We must have $q = 2, 3$ or 4.

Following our paper referred to in the introduction, we shall call the sum of the orders of the branch points of an algebraic function at a given point the *index* of the function at that point. The sum of the indices of the inverse of a rational function of degree n is $2n - 2$.

If, when we represent the substitution at a critical point in the form (2), the coefficient a belongs to the exponent d modulo n, where $d > 1$, the substitution is of order d, and the index at the critical point is $(n - 1)(d - 1)/d$.

We shall consider first those cases in which the Riemann surface for w has a branch point of order $n - 1$. Such a branch point must be present at infinity if w is the inverse of a polynomial. As the index of w at any critical point is at least $(n - 1)/2$, and as the sum of the indices of w is $2n - 2$, we must have, in this case, $q = 2$ or $q = 3$.

If $q = 2$, w must have two branch points of order $n - 1$. Subjecting z to a suitable linear transformation, we may suppose that one of these points is at infinity and the other at zero. The surface thus obtained is recognized as that for $w = z^{1/n}$. The functions uniform on it are rational functions of $z^{1/n}$. Of these, the only ones which are inverses of polynomials are linear integral functions of $z^{1/n}$.

If $q = 3$, the remaining two critical points of w, since each has an index not less than $(n - 1)/2$, must each have precisely $(n - 1)/2$ as index. Subjecting z to a suitable linear transformation, we may suppose that the branch point of order $n - 1$ is at infinity, and the other two critical points at $z = 1$ and $z = -1$ respectively. The substitutions at the latter points are of the form

$$\nu' \equiv -\nu + h_1, \qquad \nu' \equiv -\nu + h_2 \pmod{n},$$

respectively. The substitution at infinity, the result of following the second of these two by the first, is

$$\nu' \equiv \nu + h_1 - h_2 \pmod{n}.$$

As this substitution is of period n, we have $h_1 \not\equiv h_2$. If we renumber the branches of w, giving to the branch numbered ν the number μ determined by the congruence

$$\nu \equiv (h_1 - h_2)\mu + \frac{n+1}{2} h_1 \pmod{n}, \tag{5}$$

the three elementary substitutions become

$$\mu' \equiv -\mu, \qquad \mu' \equiv -\mu - 1, \qquad \mu' \equiv \mu + 1 \pmod{n}.$$

Thus only one mechanism is possible for the surface of w. Now it is well known that the trigonometric polynomial of degree n, $f_n(w)$, defined by the relation

$$\cos nu = f_n(\cos u)$$

has an inverse expressible in terms of radicals, the critical points of the inverse being at 1, -1, and ∞. Hence w must be a rational function of $f_n^{-1}(z)$.

Summarizing the foregoing results, we see that *the only polynomials of prime degree whose inverses can be expressed in terms of radicals are those of the forms* $a(w+b)^n + c$ *and* $af_n(bw+c) + d$, *where* $\cos nu = f_n(\cos u)$.

We consider now the case of $q = 4$. As the index of w at each critical point is at least $(n-1)/2$ and as the sum of the indices is $2n-2$, the index of w is precisely $(n-1)/2$ at each critical point. The corresponding substitutions are all of order 2. Subjecting z to a suitable linear transformation, we may throw one of the critical points to ∞, and so dispose the others that the sum of their affixes e_1, e_2, and e_3 is zero. Let the substitutions at e_1, e_2, and e_3 be respectively

$$\nu' \equiv -\nu + h_1, \qquad \nu' \equiv -\nu + h_2, \qquad \nu' \equiv -\nu + h_3 \pmod{n}.$$

Suppose that h_1 and h_2 are unequal. If we give to the branch of w numbered ν the number μ determined by (5), the three substitutions become

$$\mu' \equiv -\mu, \qquad \mu' \equiv -\mu - 1,$$
$$(h_1 - h_2)\mu' \equiv (h_2 - h_1)\mu + h_3 - h_1 \pmod{n}.$$

As h_3 varies from 0 to $n-1$, we obtain n types of surfaces which, it will be seen below, are all distinct.

If h_1 and h_2 are equal, they must be distinct from h_3, else the surface would not hang together. We can in this case reduce the three substitutions to

$$\mu' \equiv -\mu, \qquad \mu' \equiv -\mu, \qquad \mu' \equiv -\mu - 1 \pmod{n}.$$

Thus, the critical points being disposed as described above, there are at most $n+1$ distinct surfaces for w. To identify these, we construct the elliptic function $\wp(u|\omega_1, \omega_3)$, with $\wp(\omega_i) = e_i$ $(i = 1, 2, 3)$. This is possible, since $e_1 + e_2 + e_3 = 0$. Let

$$\begin{aligned} \Omega_1 &= a\omega_1 + b\omega_3, \\ \Omega_3 &= c\omega_1 + d\omega_3, \end{aligned} \tag{6}$$

where $ad - bc = n$. It is well known that there are $n+1$ distinct transformations (6), and that for every transformation, we have

$$\wp(u|\omega_1, \omega_3) = R[\wp(u|\Omega_1, \Omega_3)],$$

where $R(w)$ is a fractional rational function of degree n, whose inverse can

be expressed in terms of radicals. The critical points of $R^{-1}(z)$ will correspond to those values which $\wp(u \mid \omega_1, \omega_3)$ assumes twice at a point, namely, e_1, e_2, e_3, and ∞. It is easy to show that the $n+1$ Riemann surfaces for the inverses of the functions $R(w)$, occurring in the $n+1$ distinct transformations (6), have distinct mechanisms. The $n+1$ surfaces exhibited above can be none other than these.

We pass finally to the case in which $q = 3$ and in which no branch point of order $n-1$ exists. We must have, by (3),

$$\frac{1}{s_1} + \frac{1}{s_2} + \frac{1}{s_3} = 1.$$

There are the three possibilities:

(*a*) $\qquad s_1 = s_2 = \tfrac{1}{4}, \qquad s_3 = \tfrac{1}{2};$

(*b*) $\qquad s_1 = s_2 = s_3 = \tfrac{1}{3};$

(*c*) $\qquad s_1 = \tfrac{1}{2}, \qquad s_2 = \tfrac{1}{3}, \qquad s_3 = \tfrac{1}{6}.$

Consider Case (*a*). As every s_i is a divisor of $n-1$, we must have $n \equiv 1 \pmod 4$. Subjecting z to a suitable transformation, we can place the critical points with substitutions of order 4 at 0 and ∞ and the third critical point at any point e_1^2. Normalizing the substitutions at the critical points, we find that there are not more than two distinct mechanisms for the surface. We now take $\wp u$ so that $\wp(\omega_1) = e_1$, $\wp(\omega_2) = 0$, and $\wp(\omega_3) = -e_1$. This corresponds to the lemniscatic case. Now, as $n \equiv 1 \pmod 4$, we have $n = \alpha^2 + \beta^2$, where α and β are integers. In the lemniscatic case, the two functions $\wp^2(\alpha \pm \beta i)u$ are rational functions of $\wp^2 u$.* The rational functions thus obtained can be inverted in terms of radicals. It is not hard to identify their surfaces with those of Case (*a*).

Similarly, it is found that Cases (*b*) and (*c*) lead to the rational functions mentioned in Case (*e*) of the introduction, which are met in the multiplication formulas for the equianharmonic case.†

III. Polynomials of composite degree

We consider a polynomial $w = F(z)$, of prime or composite degree n, and seek those cases in which $F^{-1}(w)$, the inverse of $F(z)$, can be expressed in terms of radicals.‡

* For details relative to complex multiplication in the lemniscatic and equianharmonic cases, see Ritt, *Periodic functions with a multiplication theorem*, these Transactions, vol. 23 (1922), p. 16.

† A detailed discussion of all Riemann surfaces considered in this section is contained in a paper by the writer, *Permutable rational functions*, now in the hands of the editors of these Transactions.

‡ We have interchanged the rôles of w and z in order to conform with the notation of our paper referred to above.

If $F(z)$ is a composite polynomial, that is, if

$$F(z) = \phi_1[\phi_2(z)],$$

where $\phi_1(z)$ and $\phi_2(z)$ are of degrees greater than unity, then if $F^{-1}(w)$ can be expressed in terms of radicals, $\phi_1^{-1}(w)$ and $\phi_2^{-1}(w)$ can also be so expressed, for

$$\phi_1^{-1}(w) = \phi_2[F^{-1}(w)], \qquad \phi_2^{-1}(w) = F^{-1}[\phi_1(w)].$$

We may thus restrict ourselves to the determination of prime polynomials whose inverses can be expressed in terms of radicals.

We will show that if the inverse of a *prime* polynomial is expressible in terms of radicals, the degree of the polynomial, if not equal to four, is a prime number. Thus, using the result found for polynomials in the preceding section, we will know that *if* $F(z)$ *has the decomposition into prime polynomials*

$$F = \phi_1 \phi_2 \cdots \phi_r$$

each $\phi_i(z)$ *is either of degree* 4, *or else is of the form* $\lambda_1 \pi\lambda_2$, *where* $\lambda_1(z)$ *and* $\lambda_2(z)$ *are linear and where* $\pi(z)$ *is either a prime power of* z *or a trigonometric polynomial of prime degree.**

We refer to § II of our paper, *Prime and composite polynomials*, for a proof of the fact that *a necessary and sufficient condition that a polynomial be prime is that the group of monodromy of its inverse be primitive.*

It is well known that the degree of a primitive solvable group is a power of a prime. As the substitution corresponding to the branch point at infinity of the inverse of a polynomial of degree n consists of a single cycle of n letters, the proof that prime polynomials whose inverses can be expressed in terms of radicals are either of prime degree or of degree 4 will be complete as soon as we have proved the following theorem on substitution groups:

THEOREM. *A primitive solvable group in* p^m *letters with* p *prime and* $m > 1$ *cannot contain a substitution of order* p^m, *except in the case of* $p = 2$, $m = 2$.

Let G be a primitive solvable group of degree p^m. Suppose that G contains a cyclic subgroup C of order p^m.

It is well known that G contains an invariant transitive abelian subgroup Γ, of order p^m, every substitution of which, except identity, is of order p. As Γ is permutable with C, these two groups generate a group H in G, the order of which is the product of the orders of Γ and C, p^{2m}, divided by the order of the group of substitutions common to Γ and C. The order of H is a power of p greater than p^m. The substitutions of H are all of the form $c\gamma$ where c and γ are substitutions of C and Γ respectively.

The group C is invariant in a subgroup of H of order greater than p^m.

*Certainly if each $\varphi_i(z)$ is of one of the three types described, $F^{-1}(w)$ can be expressed in terms of radicals.

Hence there must be substitutions of H which are not in C, and with respect to which C is invariant. Suppose that $c\gamma$ is such a substitution, where γ is not in C. Then, since

$$\gamma^{-1} c^{-1} C c \gamma = \gamma^{-1} C \gamma = C,$$

C is invariant with respect to certain substitutions γ of Γ, which are not in C.

Let

$$c_1 = (0\ 1\ 2 \cdots \nu \cdots p^m - 1)$$

be the substitution which generates C. We shall determine the group of substitutions which converts c_1 into a power of itself. Let α be a substitution such that

$$\alpha^{-1} c_1 \alpha = c_1^r,$$

where r is any integer not divisible by p. It is evident that α is determined as soon as r, and the index s by which α replaces 0, are given. Consider the substitution given analytically by

$$\nu' \equiv r\nu + s \pmod{p^m}. \tag{7}$$

Noting that c_1 has the representation $\nu' \equiv \nu + 1 \pmod{p^m}$, we see that $\alpha^{-1} c_1 \alpha$ has the representation $\nu' \equiv \nu + r \pmod{p^m}$. Thus α transforms c_1 into c_1^r and replaces 0 by s. The group in which C is invariant is given by (7), where r assumes all values prime to p, and where s is unrestricted.

We shall now impose the condition that a substitution of the form (7) belong to Γ, but not to C. We have for α^p the representation

$$\nu' \equiv r^p \nu + s(r^{p-1} + r^{p-2} + \cdots + 1) \pmod{p^m}.$$

As α is not in C, we cannot have $r \equiv 1 \pmod{p^m}$. Since α belongs to Γ, α^p is identity, so that

$$r^p \equiv 1 \pmod{p^m}, \tag{8}$$

$$s(r^{p-1} + r^{p-2} + \cdots + 1) \equiv 0 \pmod{p^m}. \tag{9}$$

By Fermat's theorem,

$$r^p \equiv r \pmod{p}, \tag{10}$$

so that, by (8) and (10), $r \equiv 1 \pmod{p}$. Let $r = kp + 1$. We have

$$r^{p-i} \equiv (p - i)kp + 1 \pmod{p^2},$$

so that

$$r^{p-1} + r^{p-2} + \cdots + 1 \equiv \sum_{i=1}^{i=p} [(p - i)kp + 1] \pmod{p^2},$$

$$\equiv \frac{kp^2(p-1)}{2} + p \pmod{p^2}.$$

Suppose that $p > 2$. Then $p - 1$ is even and

$$r^{p-1} + r^{p-2} + \cdots + 1 \equiv p \pmod{p^2}. \tag{11}$$

That is, the first member of (11) is divisible by p, but not by p^2. Hence, referring to (9), we see that s is divisible by p^{m-1}.

Suppose then that α has the form

$$\nu' \equiv (kp + 1)\nu + lp^{m-1} \pmod{p^m},$$

where kp is not divisible by p^m. Since Γ is regular, if we can show that α leaves certain indices fixed, we will know that α cannot belong to Γ. Consider the congruence

$$(kp + 1)\nu + lp^{m-1} \equiv \nu \pmod{p^m},$$

or

$$kp\nu + lp^{m-1} \equiv 0 \pmod{p^m}. \tag{12}$$

Since kp is not divisible by p^m, we see that (12) must have roots.

We conclude that for α to belong to Γ and not to C, we must have $p = 2$. If $p = 2$, we must have, by (8),

$$r^2 \equiv 1 \pmod{2^m}.$$

If $m > 2$, this congruence has the four solutions

$$r \equiv \pm 1, \qquad r \equiv 2^{m-1} \pm 1 \pmod{2^m}.$$

Suppose that $m > 2$, and that $r \equiv 2^{m-1} + 1 \pmod{2^m}$. The highest power of 2 by which $r + 1$ is divisible is the first. Since, by (9), $s(r + 1)$ is divisible by 2^m, we see that s is divisible by 2^{m-1}. That is, α has the form

$$\nu' \equiv (2^{m-1} + 1)\nu + 2^{m-1} l \pmod{2^m}.$$

Since this substitution leaves the index 0, or the index 1, fixed, according as l is even or odd, we cannot have $r \equiv 2^{m-1} + 1 \pmod{2^m}$, if $m > 2$.

Suppose that $m > 2$, and that $r \equiv 2^{m-1} - 1 \pmod{2^m}$. Then, by (9), s must be even. Putting $\nu' = \nu$, we have

$$(2^{m-1} - 2)\nu + 2l \equiv 0 \pmod{2^m}.$$

Now, since $m > 2$, $2^{m-1} - 2$ is divisible by no higher power of 2 than the first, so that the congruence above has roots, and α cannot belong to Γ.

We consider finally the case of $r \equiv -1 \pmod{2^m}$. We have the 2^m substitutions

$$\nu' \equiv -\nu + s \pmod{2^m}, \tag{13}$$

which transform C into itself. If s is even, (13) leaves two letters fixed and cannot belong to Γ. Consider those substitutions for which s is odd. We say that if Γ contains one of them, it contains all of them. For if,

in the substitutions

$$\nu' \equiv -\nu + s_1, \qquad \nu' \equiv -\nu + s_2 \pmod{2^m},$$

s_1 and s_2 are both odd, the substitution $\nu' \equiv \nu + (s_2 - s_1)/2 \pmod{2^m}$, which belongs to C', transforms the first into the second. As Γ is an invariant subgroup, if either of these substitutions belongs to Γ the other does also.

Suppose that Γ contains the two substitutions

$$\nu' \equiv -\nu + 1, \qquad \nu' \equiv -\nu - 1 \pmod{2^m},$$

which we denote by α_1 and α_2 respectively. As Γ is abelian, we have, equating $\alpha_1\alpha_2$ and $\alpha_2\alpha_1$,

$$\nu - 2 \equiv \nu + 2 \pmod{2^m},$$

from which it follows that $m = 2$.

In the case of $p = 2$, $m = 2$, the symmetric group in four letters has substitutions of order 4.

The proof of the theorem is completed.

Columbia University,
New York, N. Y.

ON THE LOCATION OF THE ROOTS OF THE JACOBIAN OF TWO BINARY FORMS, AND OF THE DERIVATIVE OF A RATIONAL FUNCTION*

BY

J. L. WALSH

1. **Introduction.** There have recently been published the following results:†

THEOREM I. *If the points z_1, z_2, z_3 vary independently and have circular regions as their respective loci, then the locus of the point z_4 defined by the real constant cross ratio*

$$\lambda = (z_1, z_2, z_3, z_4)$$

is also a circular region.

THEOREM II. *Let f_1 and f_2 be binary forms of degrees p_1 and p_2 respectively, and let the circular regions C_1, C_2, C_3 be the respective loci of m roots of f_1, the remaining $p_1 - m$ roots of f_1, and all the roots of f_2. Denote by C_4 the circular region which is the locus of points z_4 such that*

$$(z_1, z_2, z_3, z_4) = \frac{p_1}{m},$$

when z_1, z_2, z_3 have the respective loci C_1, C_2, C_3. Then the locus of the roots of the jacobian of f_1 and f_2 is composed of the region C_4 together with the regions C_1, C_2, C_3, except that among the latter the corresponding region is to be omitted‡ if any of the numbers m, $p_1 - m$, p_2 is unity. If a region C_i $(i = 1, 2, 3, 4)$ has no point in common with any other of those regions which is a part of the locus of the roots of the jacobian, it contains precisely $m - 1$, $p_1 - m - 1$, $p_2 - 1$, or 1 of those roots according as $i = 1, 2, 3,$ or 4.

It is the primary object of the present paper to consider extensions of Theorem II in various directions. Chapter I studies the possibility of extend-

* Presented to the Society, September 9, 1921.

† These Transactions, vol. 22 (1921), pp. 101–116; this paper will be referred to as II. It was preceded by a paper, these Transactions, vol. 19 (1918), pp. 291–298, which will be referred to as I, and was followed by a third paper, these Transactions, vol. 23 (1922), pp. 67–88, which will be referred to as III. We shall also have occasion to refer to two other of our papers, using the letters A and S respectively: Annals of Mathematics (2), vol. 22 (1920), pp. 128–144; *Comptes Rendus du Congrès International des Mathématiciens*, Strasbourg, 1920, pp. 349–352. The term *locus* in Theorem I of the present paper replaces the term *envelope* used in II.

‡ The corresponding region is to be omitted in this *enumeration* of the points of the locus; the corresponding region may nevertheless be in whole or in part, a portion of the locus of the roots of the jacobian.

ing Theorem II to include regions which are not circular. It is found (Theorem III) that the reasoning formerly used cannot be directly extended, and specific examples bring out the nature of the difficulties in supplying other modes of reasoning. Chapter II treats the extension of Theorem II by increasing the number of circular regions which are allowed to be loci of roots of the ground forms of the jacobian. Theorems VI and XI are fairly general results obtained by this extension, the principal results of the paper. Chapter III is a short chapter which deals with centers of gravity; the results are mainly generalizations of well known results for polynomials and their derivatives. Finally, Chapter IV deals with the case of the roots of the jacobian of real forms. Theorem XV is a rather general result which applies to the roots of the derivative of a polynomial which has only real roots.

Chapter I: On the extension of Theorem II to other than circular regions

2. **A distinctive property of circular regions.** We shall now undertake to consider the extension of Theorem II to regions which are not circular. Inasmuch as rather large extensions of Theorem I in this direction can be obtained without difficulty, and in fact have been obtained in III, our first tendency is to attempt to extend Theorem II by repeating our previous reasoning of II, p. 113. This turns out to be impossible, due to the failure of Lemma I (II, p. 102) to admit of large extension to other than circular regions:

Theorem III. *If a closed region C has the property that the force at any external point P due to every set of k unit particles in C is equivalent to the force at P due to k unit particles coinciding at some point of C, then C is a circular region.*

In proving this theorem we do not need to assume the property stated for every k, but merely for any one particular k (except of course $k = 1$, for which the result is absurd). We shall suppose $k = 2$ and leave to the reader the modifications for the other values of k.

If C is the whole plane there is no external point P and we may consider the theorem true, since C is a circular region. Similarly, if C is a single point the theorem is true. In the sequel we assume C to be neither the entire plane nor a single point.

It will be noted that the force at P due to equal particles at two points M and N is equivalent to the force at P due to two coincident particles situated at Q, the harmonic conjugate of P with respect to M and N,* which situation is invariant under linear transformation. We shall proceed to prove the

Lemma. *If P is exterior to C, and M and N are two points of C, then C contains every point on that arc of the circle MNP bounded by M and N which does not contain P.*

* This fact is quite easy to prove; see for example A, pp. 128–129.

Choose P at infinity and equal particles at M and N; the point Q which is the mid point of the line segment MN is seen to be in C. Then the mid points of the segments MQ, QN are also in C, and in fact we have a set of points in C everywhere dense on the segment MN. Hence this entire segment belongs to C.

The region C contains at least two points. It follows from the lemma that C contains an infinity of points. Transform one point R of the boundary of C to infinity, and consider two other distinct points V and W of that boundary. We proceed to prove that every point of the segment VW is a point of C. We can find a sequence of points not belonging to C but approaching R. Hence there is a sequence of points (the harmonic conjugates of the former sequence with respect to V and W) belonging to C and approaching the mid point U of VW, so U belongs to C. Then the mid points of UV and VW also belong to C, and in this way we prove that every point of VW belongs to C. But R also belongs to C, and hence we can prove that either every point of that infinite segment of RV which does not contain W belongs to C or every point of that infinite segment RW which does not contain V belongs to C; for definiteness suppose the latter.

There exists an infinite sequence of points not belonging to C but approaching W, and we assume as of course we may do that we have oriented the line VW horizontally and that we have an infinite sequence of such points $\{Z_k\}$ all lying in the lower half plane. If a line through Z_k cuts the segment VWR, then every point of that line which lies in the upper half plane belongs to C, by the lemma. For any preassigned point Y in the upper half plane we can choose a point Z_k such that YZ_k cuts the segment VWR, so every point of the upper half plane belongs to C as does also every point of the line VW.

If the region C does not consist of precisely the upper half plane including its boundary, there is a point X of C in the lower half plane. We can make use of the fact that R is a point of the boundary of C as before, and prove that the entire finite segment joining X to an arbitrary point of the line VW belongs to C. Then V and W are not boundary points of C, contrary to our assumption. The demonstration of the theorem is now complete.

Theorem III is quite easily proved, although by essentially the same methods, if we assume C to be bounded by a regular curve. Thus the property considered is invariant under linear transformation, for the position of the k coincident particles is uniquely determined by P and the original particles, and by Theorem I (I, p. 291 = II, p. 101) P is a root of the jacobian of the two binary forms each of degree k and whose roots are respectively the k original particles in C and the k coincident particles. If C is not bounded by a circle, its boundary can be transformed into a contour which is not convex.* Then

* See Theorem III of a note by the present writer, Annals of Mathematics (2), vol. 22 (1921), pp. 262–266.

there are two points A and B on the boundary of C such that the segment AB is exterior to C and such that there is a point P exterior to C and on the line AB but not on the segment AB. The force at P due to one particle at A and $k-1$ particles at B is equal to the force at P due to k coincident particles at a point which is on the segment AB but which coincides with neither A nor B and hence which is exterior to C.

The essence of Lemma I (II, p. 102) was proved and applied by Laguerre;* his formulation of the result was quite different from the present formulation, although the application was to the location of the roots of algebraic equations. Thus Theorem III appears as a sort of converse of the theorem of Laguerre, as well as of Lemma I (II, p. 102).

3. **Successive application of Theorem II to the determination of loci.** The property of circular regions stated in Theorem III seems to be conclusive in showing that the reasoning of II, p. 113, cannot be reproduced to give large extensions of Theorem II. Theorem III justifies, moreover, the somewhat artificial use of circular regions in III, Theorem XIII.

We now point out by two simple examples how results can be found from successive applications of Theorem II. These examples are not given as large extensions of Theorem II, but rather to show how difficult is that extension to regions which are not circular. The proofs of these theorems are left to the reader; the proof of the latter depends on III, Theorem IX.

THEOREM IV. *Suppose we have a finite or infinite number of sets of regions* $C_1^{(n)}$, $C_2^{(n)}$, $C_3^{(n)}$, $C_4^{(n)}$, *of Theorem* I *corresponding to the value* $\lambda = p_1/m$, *and suppose that no* $C_i^{(n)}$ *has a point in common with* $C_j^{(k)}$ *unless* $i = j$ $(i, j = 1, 2, 3, 4)$. *Denote by* T_1, T_2, T_3, T_4, *the regions common to all the* $C_1^{(n)}$, $C_2^{(n)}$, $C_3^{(n)}$, $C_4^{(n)}$, *respectively. Then if* T_1 *contains* m *roots of a bilinear form* f_1, *if* T_2 *contains all the remaining* $p_1 - m$ *roots of* f_1, *and if* T_3 *contains all the* p_2 *roots of a second form* f_2, *then the regions* T_1, T_2, T_3, T_4 *contain all the roots of the jacobian of* f_1 *and* f_2. *No two of the regions* T_1, T_2, T_3, T_4 *have a point in common, and they contain respectively* $m-1$, $p_1 - m - 1$, $p_2 - 1$, 1 *of the roots of the jacobian.*

There is no reason to suppose that the actual locus of the roots of the jacobian is composed of T_1, T_2, T_3, T_4, when T_1, T_2, T_3 are the loci of roots of the ground forms. But if the regions $C_1^{(n)}$, $C_2^{(n)}$, $C_3^{(n)}$, $C_4^{(n)}$ have the disposition suggested in the first part of § 11 (III) or more generally if T_4 is the locus of the point z_4 determined by its cross ratio p_1/m with the points z_1, z_2, z_3 whose loci are T_1, T_2, T_3, these four regions form that locus, except of course that among these latter three the corresponding region is to be omitted if any of the numbers m, $p_1 - m$, p_2 is unity.

*(Œuvres, pp. 56–63; p. 59.

THEOREM V. *In the situation of* III, *Theorem* VIII, *suppose the point* P *which is a center of external similitude for every pair of the circles* C_1, C_2, C *to be actually external to all those circles. Denote by* T_1, T_2, T *the portions of the interiors of those circles which lie between two half lines through* P *cutting those circles. Then if* T_1 *and* T_2 *are the respective loci of* m_1 *and* m_2 *roots of a polynomial* $f(z)$, *the regions* T_1, T_2, T *are the loci of the roots of its derivative* $f'(z)$, *except that* T_1 *or* T_2 *is to be omitted if* m_1 *or* m_2 *is unity. If* T_1, T_2, T *are mutually external, they contain respectively* $m_1 - 1$, $m_2 - 1$, 1 *of the roots of* $f'(z)$.

A theorem similar to Theorem V will be obtained by cutting the circles C_1, C_2, C by any convex contour, but no result can generally be stated in this case concerning the actual locus of the roots of $f'(z)$.

In discussing the possibility of the extension of Theorem II by reproducing the reasoning of II, p. 113, we reached the impossibility of extending Lemma I (II, p. 102), by which we may replace the force at an arbitrary point P exterior to a region C due to k particles in C by the force at P due to k coincident particles in C. To obtain certain facts, however, concerning the location of the roots of the jacobian, it may not be necessary to replace the k particles in C by k coincident particles in C *for an arbitrary point* P exterior to C but merely *for certain points* P exterior to C. This fact is notably true if the ground forms are real, as we shall show in Chapter IV; it is also true for certain other cases, as we shall now indicate by an extremely simple example.

We consider the polynomial

$$f(z) = z(z - \alpha_1)(z - \alpha_2),$$

where α_1 and α_2 have as their common locus the interior and boundary of the circle C_1 whose center is $z = 6$ and radius unity. Then $f'(z)$ has a root z_1 which has as its locus the interior and boundary of the circle C_1 and a root z_2 which has as its locus the interior and boundary of the circle C whose center is $z = 2$ and radius $1/3$. Under the given conditions, moreover, z_1 and z_2 remain separate and distinct.

Let us now consider the same polynomial but assign to the roots α_1 and α_2 as their common locus the right-hand semicircular region S_1' of C_1. Any point of the right-hand semicircular region S' of C is by III, Theorem IX, a point of the locus of z_2, and we shall prove that no other point is a point of this locus. Suppose a point $\bar{z}$ to be a point of the locus of z_2; $\bar{z}$ is evidently in or on C'. The force at $\bar{z}$ due to particles at α_1 and α_2 is equivalent to the force at $\bar{z}$ due to two particles coinciding at some point α. We know that $\bar{z}$ is in or on C, hence exterior to C_1, so α is in or on C_1. Moreover, $\bar{z}$ is not in the half plane bounded by and lying to the right of the line $x = 6$, and hence α is in that half plane. Then α is in S_1', so $\bar{z}$ is in S'.

The locus of z_1 under the conditions stated includes of course S_1', but other points as well. In fact, if we choose $\alpha_1 = 6 + i$, $\alpha_2 = 6 - i$, the two roots of the real polynomial $f'(z)$ cannot be conjugate imaginary, so z_1 is real but to the left of the point $z = 6$ and therefore not a point of S_1'.

We consider anew the same polynomial and assign the left-hand semicircle S_1'' of C_1 as the locus of α_1 and α_2. Any point of S_1'' is evidently a point of the locus of z_1, and no other point belongs to this locus. For any such point would lie in the semicircle S', which is impossible, according to the theorem of Lucas. Under our conditions the locus of z_2 includes the left-hand semicircle S'' of C, but also other points. We choose as before $\alpha_1 = 6 + i$, $\alpha_2 = 6 - i$, and find z_2 to be real and on or within S' as previously noted. But $z_2 \neq 2$, and hence is not a point of S''.

Results which are large extensions of Theorem II to other than circular regions seem difficult to prove, as is shown by Theorems IV and V, even when the regions involved are common to two or more circular regions. But theorems of a certain type are easily established; we give simply one example:

Let the intersecting circles C_1 and C_2 with centers at the points α_1 and α_2 and radii r_1 and r_2 be the respective loci of m_1 and m_2 roots of a polynomial $f(z)$. Let the region common to C_1 and C_2 be the locus of $f(z)$, which is supposed to have no roots other than those mentioned. Then the roots of $f'(z)$ lie in C_1, C_2, and the region common to the two circles whose centers are the points

$$\frac{(m_2 + m_3)\alpha_1 + m_1\alpha_2}{m_1 + m_2 + m_3}, \qquad \frac{m_2\alpha_1 + (m_1 + m_3)\alpha_2}{m_1 + m_2 + m_3},$$

and whose radii are respectively

$$\frac{(m_2 + m_3)r_1 + m_1 r_2}{m_1 + m_2 + m_3}, \qquad \frac{m_2 r_1 + (m_1 + m_3)r_2}{m_1 + m_2 + m_3}.$$

The region mentioned in this theorem is not the *locus* of the roots of $f'(z)$, for the intersection of the last two circles mentioned in the theorem can never be a root of $f'(z)$. The proof of these facts is left to the reader.

4. **The number of roots of the jacobian in a circular region.** A question concerning the distribution of the roots of the jacobian which is very closely connected with the question of extending Theorem II to regions other than circular is that of the number of roots of the jacobian in the regions C_i when two or more of those regions have common points. Thus it might be supposed that C_1, C_2, C_3, C_4 contain *always* $m, p_1 - m, p_2, 1$ of the roots of the jacobian except for the possibility that this number be exceeded when C_i has one or more points in common with another of the regions. This supposition is false, however, as we proceed to show by an example. Thus consider the circles C_1 and C_2 in the form of Theorem II corresponding to III, Theorem VIII

(i.e., S, Theorem I), such that C intersects C_2 but is exterior to C_1. Let the line of centers of the circles intersect C_1 in A_1 (the intersection nearest C and C_2) and intersect C_2 in A_2 and C in A (the intersections farthest from C_1). When m_1 roots of $f(z)$ coincide at A_1 and m_2 roots at A_2, we have $m_1 - 1$ roots of $f'(z)$ at A_1, $m_2 - 1$ roots at A_2, and one root at A. When the m_2 roots at A_2 move slightly so that they do not all coincide but all remain on C_2 and symmetric with respect to the line A_1A_2, the force corresponding at A and in the neighborhood of A becomes equivalent to the force due to m_2 particles coinciding *exterior* to C_2, since these particles can be considered to lie in the circular region consisting of the circle C_2 and all exterior points. Hence the root of $f'(z)$ at A moves and becomes exterior to C; the m_2 roots at A_2 remain in the vicinity of A_2 and there is no root of $f'(z)$ on or interior to C.

The question which we raised has thus been answered so far as concerns a region C which contains no roots of the ground forms. The result is essentially the same for a region which does contain a number of roots of the ground forms. Consider the case of the derivative of a polynomial, the second theorem on page 115 of II, locate m_1 roots of $f(z)$ at the null circle C_1, and locate the remaining m_2 roots at two points A and B different from C_1. There are but two roots z_1 and z_2 of $f'(z)$ distinct from A, B, and C_1, and these are interior to the triangle ABC_1, so a circle C_2 can be drawn which includes A and B (that is, m_2 roots of $f(z)$) but includes neither z_1 nor z_2 and hence contains only $m_2 - 2$ roots of $f'(z)$.

The question we have been considering is closely connected with the following:* Suppose a circle C contains at least r roots of a polynomial $f(z)$ of degree n. What can be said of the number of roots of $f'(z)$ in C?

On the one hand, C may contain all the roots of $f(z)$ and hence all the $n - 1$ roots of $f'(z)$. On the other hand, C may contain r roots of $f(z)$ and yet no root of $f'(z)$ if merely $r < n$. In fact, we prove that C may contain precisely $n - 1$ roots of $f(z)$ and contain a preassigned number p of the roots of $f'(z)$, if merely $p < n - 1$. Locate one root of $f(z)$ at a point P and the other $n - 1$ roots at $n - p - 1$ distinct points which lie on a line L not passing through P. Then we can describe a circle C which includes the $n - p - 1$ distinct points on L and hence p roots of $f'(z)$ but which contains no other roots of $f'(z)$.

* Still another allied question is: Suppose a circle C is known to contain at least r roots of a polynomial of degree n; determine the smallest (concentric) circle C' which always contains at least m roots of the derived polynomial.

The circle C' exists only if $m < r$. For $n = r = m + 1$, the answer is given by Lucas's Theorem. For the case $r = 2$, the circle C' is readily determined by means of a theorem due to Grace, to which reference is made in A, § 4. For the case $n - 1 = r = m + 1$, the circle C' is easily found by the second theorem of II, p. 115. For other cases the problem seems considerably more complicated.

Chapter II: On the extension of Theorem II to a larger number of circular regions

5. **Problem of the locus corresponding to any number of circular regions.** Our attempt in Chapter I to extend Theorem II in a form to apply to the jacobian of two particular binary forms by considering regions other than circular as loci of the roots of the ground forms and finding the corresponding locus of the roots of the jacobian was not particularly fruitful. This seems to result rather from the nature of the problem itself than from the precise methods employed. We now take up the possibility of extending Theorem II so as to consider not merely three circular regions but any number of circular regions. Let us suppose explicitly that we have the binary forms f_1 and f_2 of respective degrees p_1 and p_2, and that the circular regions C'_1, C'_2, $\cdots$, C'_m are the loci respectively of p'_1, p'_2, $\cdots$, p'_m roots of f_1 and the circular regions C''_1, C''_2, $\cdots$, C''_n are the loci respectively of p''_1, p''_2, $\cdots$, p''_n roots of f_2, where we have

$$p'_1 + p'_2 + \cdots + p'_m = p_1,$$
$$p''_1 + p''_2 + \cdots + p''_n = p_2.$$

For convenience in phraseology, we shall suppose that none of these regions is either a point or the entire plane unless otherwise stated. We wish then to find the location of the roots of the jacobian of f_1 and f_2; not merely to determine certain regions in which lie or do not lie the roots of the jacobian, but to determine the actual *locus* of those roots under the assigned conditions, as in Theorem II.

Let us consider, for any particular values of the roots of f_1 and f_2 which satisfy our hypothesis, a root ζ of the jacobian exterior to all the circular regions C'_1, $\cdots$, C'_m, C''_1, $\cdots$, C''_n. This root ζ is an analytic function of α, any root of f_1 or f_2, and hence when α varies over a certain two-dimensional continuum, ζ also varies over a certain two-dimensional continuum. We thus have a certain number of regions which may or may not be distinct and may or may not have common points which are the loci of the points ζ. We see by the analyticity of the transformation that all the points α must be on the boundaries of their proper regions whenever a point ζ corresponding is on the boundary of its locus.* Moreover, if a point ζ is on the boundary of its locus, and exterior to all the regions C'_1, $\cdots$, C'_m, C''_1, $\cdots$, C''_n, we know by Lemma I (II, p. 102) that all the points α pertaining to any one circular region can be considered to coincide on the boundary of that region. But the precise manner of simultaneous variation of these coincident roots on the

* There is an exception to this reasoning if the algebraic equation defining ζ degenerates and if ζ is independent of a particular α, but in that case α can be moved at will without changing ζ and so α can be considered as on the boundary of its locus. A similar remark applies also below.

boundaries of their loci in such a manner that a point ζ or several points ζ remain on the boundaries of their loci and trace out those boundaries is as yet unknown.

Let us restrict ourselves for the moment to the situation where the circular regions $C_1', \cdots, C_n''$ are relatively small, or to be more precise, such that for no choice of the roots of the ground forms in their proper regions can two roots of the jacobian coalesce exterior to those circular regions; we suppose further that no two of the regions $C_1', \cdots, C_n''$ and the regions R which are the loci each of one of the roots of the jacobian exterior to those circular regions when the roots of the ground forms have their proper regions as loci—no two of all these regions have a point in common. We may allow the roots of the ground forms to coalesce in their proper regions; we notice that the circular regions $C_1', C_2', \cdots, C_m', C_1'', C_2'', \cdots, C_n''$ contain and are therefore the loci of respectively $p_1' - 1, p_2' - 1, \cdots, p_m' - 1, p_1'' - 1, p_2'' - 1, \cdots, p_n'' - 1$ roots of the jacobian. There are then $m + n - 2$ regions R each of which is the locus of one root of the jacobian. When we allow the circular regions to become larger and larger, of course the regions R expand also, need not preserve their identity (for example, two of them may coincide), and finally these regions cover the entire plane.

Very little is known of the precise nature of the boundaries of these regions R.* Their boundaries are not, except in very special instances, circular regions, but are curves which presumably have interesting properties with reference to the boundaries of the regions $C_1', \cdots, C_n''$, which properties can be expressed in a manner so as to be invariant under linear transformation. It is evidently true that if we start with any situation $C_1', \cdots, C_n''$ and if we allow two of the regions $C_1', \cdots, C_m'$ or two of the regions $C_1'', \cdots, C_n''$ to coalesce, one of the regions R will coalesce with them, and we shall have precisely the situation of $m - 1$ regions C_i' or $n - 1$ regions C_i''.

The exact determination of the regions R in any very general case seems difficult. If all the original circular regions reduce to points except one of them, say C_1', we can determine the path of the roots ζ of the jacobian as α, a p_1'-fold root of f_1, traces the circle C_1'. These roots ζ, in their totality, trace closed curves, for the situation when α returns to its initial position is exactly the same as the initial situation. The boundaries of the regions R must be composed of these closed curves, or at least of portions of them. If now we allow a second one of our circular regions, say C_1'', to be a non-degenerate region, the new locus of the roots of the jacobian will be a number of regions

* The writer conjectures that when there are q roots of the jacobian in these regions R, these regions are in their totality bounded by a degenerate or non-degenerate q-circular $2q$-ic; only the degenerate cases of this curve have ever been treated in detail, except for $q = 1$. Compare Walsh, Proceedings of the National Academy of Sciences, vol. 8 (1922), pp. 139–141.

R'. The boundaries of the regions R' will be curves which are envelopes of the curves R corresponding to the region C_1' and the null regions $C_2', \cdots, C_m', C_1'' = \beta, C_2'', \cdots, C_n''$, while the point β traces the circle C_1''. By continuing in this way, we have a process for the generation of the regions R in any case desired. But the actual determination of the boundaries in a very general case would presumably be too laborious by this process; more powerful methods will have to be devised.

The statement has been made that the regions R are not in general circular regions; it is perhaps worth while to present a specific instance to illustrate this fact. We consider the polynomial

$$f(z) = (z - \alpha_1)(z - \alpha_2)(z - \alpha_3),$$

where $\alpha_1 = i$, $\alpha_2 = -i$, and the locus of α_3 is chosen to be a circle C_3 and its exterior, whose center is the origin and radius so large that the loci of the two roots of $f'(z)$ are entirely distinct. In the field of force to determine the roots of $f'(z)$, the force at a point of either coördinate axis due to the two particles at i and $-i$ is in direction along that axis. Hence, whenever $f'(z)$ has a root on an axis, α_3 must also be on that axis. The root ζ of $f'(z)$ larger in absolute value is determined on the positive half of the axis of reals by the particle α_3 at the right-hand intersection of C_3 and that axis, and a second particle of twice the mass at the harmonic conjugate* of ζ with respect to the points i and $-i$; this harmonic conjugate lies *to the left of* the origin. This root ζ is determined on the positive half of the axis of imaginaries by α at the upper intersection of C_3 and that axis, and a second particle of twice the mass at the harmonic conjugate of ζ with respect to i and $-i$; this harmonic conjugate lies *above* the origin. The curve bounding the locus of ζ is symmetric with respect to the coördinate axes and hence is not a circle.

The characteristic of Theorem II (and indeed also of Theorems VI and XI) in comparison with the more general results indicated in this present section seems to be a certain *linearity*. This fact is brought out very clearly in S, but also in II, since by Lemma II (II, p. 102) the position of equilibrium is determined by its cross ratio with three points, a relation which is essentially linear. It is as a result of that linearity that for our particular situations the loci of the roots of the jacobian are all bounded by circles.

6. **A condition that a root of the jacobian be on the boundary of its locus.** We return now to the general case of the preceding section, and shall obtain a geometric relation between the roots of the ground forms and a root of the jacobian, when all of those roots are on the boundaries of their loci.

Consider the points α_1 and α_2 at which coincide all the p_1' roots in C_1' and all the p_2' roots in C_2' respectively; we consider also a root ζ of the jacobian which

* Compare §§ 2, 9; also A, pp. 128–129.

is supposed not to lie at a common root of the two ground forms or at a multiple root of either form, so that ζ is given by a certain algebraic equation which actually contains α_1 and α_2. We may write this equation in the form

$$k + \frac{p_1'}{\zeta - \alpha_1} + \frac{p_2'}{\zeta - \alpha_2} = 0. \tag{1}$$

We can hold fast ζ and all roots of the ground forms other than α_1 and α_2, and move α_1 and α_2 depending on each other so as to satisfy (1). Since (1) is linear in α_1 and α_2, when one of these points is made to trace a circle the other also traces a circle. When α_1 moves so as to trace C_1', α_2 moves so as to trace a circle tangent to C_2'. In fact, if α_2 were to trace a circle intersecting C_2', α_2 would at some time move *interior* to the region C_2', and still α_1 and α_2 would be in their proper loci. Then motion of α_2 holding α_1 fast would cause ζ to move over a two-dimensional continuum, so ζ would not be on the boundary of its locus.

It will be useful to study in some detail the relation between α_1 and α_2 defined by (1). The two double points of the transformation (α_1, α_2) are

$$\alpha_1 = \alpha_2 = \zeta; \qquad \alpha_1 = \alpha_2 = \zeta + \frac{p_1' + p_2'}{k};$$

denote this latter point by α. We have, of course, $p_1' + p_2' \neq 0$, so that these two points are distinct. The cross ratio of these fixed points with α_1 and α_2 in any position is readily computed:

$$\frac{[\alpha_1 - \alpha_2]\left[\zeta - \left(\zeta + \frac{p_1' + p_2'}{k}\right)\right]}{[\alpha_2 - \zeta]\left[\left(\zeta + \frac{p_1' + p_2'}{k}\right) - \alpha_2\right]} = \frac{p_1' + p_2'}{p_2'}.$$

Inasmuch as this cross ratio is real, the four points $\alpha_1, \alpha_2, \zeta, \alpha$ lie on a circle C.

The circle C is self-corresponding under the transformation (α_1, α_2). In fact, two of its points ζ and α are unchanged while a third point α_1 is transformed into a point α_2 of the circle; this is sufficient.

If p_1' and p_2' are both negative instead of both positive, we have the case of the roots of f_2 located in C_1'' and C_2''. In either of these situations, the first case we consider, the points α_1 and α_2 are separated by ζ and α. For a transformation can be made so that $k = 0$. The value of the cross ratio gives us

$$\frac{\alpha_1 - \alpha_2}{\zeta - \alpha_2} = \frac{p_1' + p_2'}{p_2'}, \qquad \frac{\alpha_1 - \zeta}{\zeta - \alpha_2} = \frac{p_1'}{p_2'};$$

so α_1 and α_2 are indeed separated by ζ and α.

We thus choose ζ as fixed on the boundary of its locus; the fixed points

α_1' and α_2' (particular values of α_1 and α_2 respectively) corresponding are also on the boundaries of their respective loci. We have already remarked that when α_1 moves from α_1' along C_1', α_2 moves from α_2' along a circle tangent to C_2'. When α_1 moves from α_1' interior to the region C_1', α_2 moves from α_2' exterior to the region C_2'. When α_1 moves on C, α_2 moves in the opposite direction but also on C. It follows that *C cuts C_1' and C_2' at angles of the same magnitude, and when C is transformed into a straight line the tangents to C_1' at α_1' and C_2' at α_2' are parallel.** There are different possibilities here according to whether C cuts C_1' and C_2' at equal angles or at supplementary angles. We leave it for the reader to notice that if the loci C_1' and C_2' are both interior or both exterior to their bounding circles, these two angles are equal; if one locus is interior and the other exterior to its bounding circle, these two angles are supplementary.

The second case we shall consider is that of two roots α_1 and α_2, of the forms f_1 and f_2, of multiplicities p_1' and p_1'', and loci C_1' and C_1'', respectively. Essentially the same formulas apply, except that in (1) and the succeeding formulas the numbers p_1' and p_2' are replaced by $p_2\,p_1'$ and $-p_1\,p_1''$ respectively; we suppose $p_2\,p_1' - p_1\,p_1'' \neq 0$. The cross ratio of the four points α_1, α_2, ζ, α shows that α_1 and α_2 are not separated by ζ or α. Hence if α_1 and α_2 trace the self-corresponding circle C, they trace it in the same sense. The angles which C cuts on C_1' and C_1'' are then equal or supplementary according as the regions C_1' and C_1'' lie one inside and the other outside their bounding circles, or as these regions lie both inside or both outside their bounding circles. When C is transformed into a straight line, the lines tangent to C_1' at α_1 and to C_1'' at α_2 are parallel.

The third case we have to treat is the remaining situation under the second case, where $p_2\,p_1' - p_1\,p_1'' = 0$; here the transformation (α_1, α_2) has but the one double point ζ. It is still true that the circle C through α_1, α_2, ζ is self-corresponding under this transformation. For if we denote by γ the point α_2 corresponding to $\alpha_1 = \alpha_2'$, where α_1' and α_2' are fixed values of α_1 and α_2, we shall have

$$-k = \frac{p_2\,p_1'}{\zeta - \alpha_1'} - \frac{p_1\,p_1''}{\zeta - \alpha_2'} = \frac{p_2\,p_1'}{\zeta - \alpha_2'} - \frac{p_1\,p_1''}{\zeta - \gamma},$$

$$\frac{(\alpha_1' - \alpha_2')(\gamma - \zeta)}{(\alpha_2' - \gamma)(\zeta - \alpha_1')} = -1,$$

so α_1', α_2', ζ, γ are concyclic; the three points α_1', α_2', ζ of C are transformed

* We cannot prove here, as in III, Theorem II, that this property holds also for the tangent to the boundary of the locus of ζ at the point ζ. In fact, if we choose another pair of points α_1', α_2', leading to the circle C', it is in general impossible for C' to cut at equal angles C_1' at α_1' and the boundary of the locus of ζ at ζ. For a specific example, see the illustration used at the close of § 5.

into three points α_2', γ, ζ of C which, therefore, is self-corresponding. If ζ is transformed to infinity, we have

$$\frac{\alpha_1' - \alpha_2'}{\alpha_2' - \gamma} = 1,$$

so γ is obtained from α_2' by translation by an amount equal to $\alpha_1' - \alpha_2'$. Then α_1 and α_2 trace C in the same sense. As in our second case, the angles which C cuts on C_1' and C_1'' are equal or supplementary according as the regions C_1' and C_1'' lie one inside and the other outside their bounding circles, or as these regions lie both inside or both outside their bounding circles. When C is transformed into a straight line, the lines tangent to C_1' at α_1 and to C_1'' at α_2 are parallel.*

We have now considered all typical cases of two roots α_1 and α_2 of the ground forms. In particular if we choose two roots of a single form which have the same locus, the reasoning we have used shows that when ζ is on the boundary of its locus α_1 and α_2 must be on the boundary of their common locus and *must coincide*. This may be used to replace Lemma I (II, p. 102). The reader may be interested in applying the remarks of the present section to the situations of Theorems II and VI.

7. **A special case of coaxial circles.** We have pointed out in § 5 that for the situation there described of $m + n$ circular regions C_1', $\cdots$, C_m', C_1'', $\cdots$, C_n'' as the loci of the roots of the ground forms, the locus of the roots of the jacobian is not in general a number of circular regions or of regions bounded by several circles. But of course there are special situations for which the locus of the roots of the jacobian is bounded by circles. This is evidently true, for example, if C_1', $\cdots$, C_m', C_1'', $\cdots$, C_n'' are bounded by $m + n$ concentric circles or more generally by $m + n$ coaxial circles having no common point.† For such a situation, moreover, the methods used in I can be applied; we shall give an illustration of that fact.

We use the notation of I, p. 293 ff., and suppose the situation simplified

* This remark enables us immediately to state something about the locus to be determined in certain cases. Thus consider the situation and notation of Theorem X (we might indeed choose that of Theorem VI). Choose the line through α_1, $\cdots$, α_n horizontal and draw a parallel L' through any point z on the boundary of its locus. If one point ζ_1 (notation of § 9) is not on L', say below L', and if z is exterior to all the circles C_1, $\cdots$, C_n, all other points ζ_i are also below L'. Then by Lucas's Theorem, z cannot be a root of $f'(z)$. Therefore all points ζ_i lie on L' and z lies on a circle C_1'.

If m_1, $\cdots$, m_n are all greater than unity, no point z interior to a circle C_i need be considered. If any m_k is unity, however, the points z interior to C_i must be considered. The writer has been unable to treat similarly (by the method just indicated) this last case, and thus completely to prove Theorem X.

† If the coaxial system is composed of all circles through two points or all circles tangent at a single point, we may consider all the roots of both forms to coincide at a single point, the jacobian vanishes identically, and the locus of the roots of the jacobian is the entire plane.

by transformation as in I. Suppose C_1 to contain k roots of f_1 (p_2 k positive particles) and l roots of f_2 (p_1 l negative particles). If l is sufficiently small in comparison with k, and if C_2 and C_3 are sufficiently remote, it seems reasonable to suppose that we can obtain a region near but exterior to C_1, which region contains no root of the jacobian. The circle C_1 contains, then, k particles each of mass p_2 and l particles each of mass $-p_1$. Outside of C_2 there are $p_1 - k$ particles each of mass p_2 and outside of C_3 there are $p_2 - l$ particles each of mass $-p_1$. If Q is a position of equilibrium, we must have, in the notation of I,

$$\frac{p_2 k}{a+r} \leqq \frac{p_1 l}{r-a} + \frac{p_2(p_1-k)}{b-r} + \frac{p_1(p_2-l)}{c+r},$$

which can be put into the equivalent form

$$\begin{aligned} 0 \leqq r^2[&- p_2 k(a+b) - p_1 l(a+c) + p_1 p_2(b+c)] \\ &+ r[p_2 k(a+b)(a-c) + p_1 l(a+c)(b-a)] \\ &+ [-p_1 p_2 a^2(b+c) + p_2 kac(a+b) + p_1 lab(a+c)]. \end{aligned} \tag{2}$$

This form does not simplify materially. Denote by C_4 and C_5 the circles whose centers are O and radii the roots of the right-hand member. The cross ratio of the points C_4', C_4'', C_5', C_5'' (notation as in I, p. 294) with the collinear points C_1'', C_2', C_3'' can easily be calculated, but this cross ratio contains a, b, c explicitly and is not independent of their ratios; we therefore use a different method to describe C_4 and C_5. We are supposing implicitly that the roots of the right-hand member of (2) are positive or that at least one of these roots is positive.

If C_4 and C_5 lie between C_1 and C_2 and between C_1 and C_3, they bound an annulus which contains no root of the jacobian. For if $r = a$, the right-hand member of (2) reduces to

$$2p_1 l(a+c)(b-a),$$

so that inequality is satisfied for $r = a$ and therefore is not satisfied when r lies between the two roots.

Under this hypothesis we can determine the precise number of roots of the jacobian in the smaller of the new circles by allowing the roots of f_1 and f_2 in C_1 to move continuously and to coincide at O. When the $p_2 k$ and $p_1 l$ particles are all in coincidence at O, the circle C_1 contains precisely $k + l - 1$ roots of the jacobian, so this is the original number of roots interior to or on the inner boundary of the annulus.

Hence we have, under the assumptions already made:

1. *If the circles C_4 and C_5 lie between C_1 and C_3, then the annular region between C_4 and C_5 contains no root of the jacobian of f_1 and f_2. The region which*

is bounded by C_4 or C_5 and contains the region C_1 contains precisely $k + l - 1$ roots of the jacobian.

2. *If C_4 and C_5 are separated by C_3, there are no roots of the jacobian in the annular region which is part of the annular region bounded by C_4 and C_5 and which contains no point of the region C_3. The circular region bounded by C_4 or C_5 which contains the region C_1 but no point of the region C_3 contains precisely $k + l - 1$ roots of the jacobian.*

This theorem can readily be expressed in general form so as to include the situation after linear transformation; compare the corresponding statement in I.

8. **Theorem VI, a general theorem for circles having an external center of similitude.** There is another fairly general class of loci other than the very simple class just considered for which the locus of the roots of the jacobian as treated in § 5 is bounded by circles. We shall now use a method which is novel in some respects but which makes use of our former results to establish

THEOREM VI. *Let the interiors and boundaries of the circles $C_1', C_2', \cdots, C_{n'}'$, whose centers are $\alpha_1', \alpha_2', \cdots, \alpha_{n'}'$, respectively, be the loci of $m_1', m_2', \cdots, m_{n'}'$ roots of the form f_1 which has no other roots. Let the interiors and boundaries of the circles $C_1'', C_2'', \cdots, C_{n''}''$, whose centers are $\alpha_1'', \alpha_2'', \cdots, \alpha_{n''}''$, respectively, be the loci of $m_1'', m_2'', \cdots, m_{n''}''$ roots of the form f_2 which has no other roots. Suppose further that a point P is interior to no circle C_i' or C_i'' and is an external center of similitude for every pair of the circles C_i', C_j' and for every pair of the circles C_i'', C_j'' and an internal center of similitude for every pair C_i', C_j''. The distinct roots of J, the jacobian of f_1 and f_2, when all the roots of the ground forms are concentrated at the centers of their proper circles, are denoted by $\alpha_1, \alpha_2, \cdots, \alpha_n$ (all of which points are collinear with P, $\alpha_1', \cdots, \alpha_{n'}', \alpha_1'', \cdots, \alpha_{n''}''$) of multiplicities $m_1, m_2, \cdots, m_n$, and by $C_1, C_2, \cdots, C_n$ are denoted the circles which have these points as centers and radii such that P is an external or internal center of similitude for every pair of the circles $C_1', \cdots, C_{n'}', C_1'', \cdots, C_{n''}'', C_1, \cdots, C_n$. Then the locus of the roots of J, the jacobian of f_1 and f_2, is composed of the interiors and boundaries of the circles $C_1, C_2, \cdots, C_n$. A circle C_i exterior to all the other circles C_j contains precisely m_i roots of J.*

Limiting cases of the circles $C_1', \cdots, C_{n'}', C_1'', \cdots, C_{n''}''$ are the points P and P', the point at infinity. We shall admit these circles as possibilities in the demonstration of the theorem, providing, however, that there is at least one proper circle, which we shall suppose to be C_1'. If there is no proper circle C_1' or C_1'', either the only roots of J are P and P', in which case the theorem remains true, or every point of the plane is a root of J, in which case the theorem is not true.

The configuration of the three sets of circles has some obvious but interesting properties relative to $\bar{J}$. Let us choose as horizontal the line L on which lie

$P, \alpha_1', \cdots, \alpha_{n'}', \alpha_1'', \cdots, \alpha_{n''}''$, with the C_i' to the right of P, let us number in order the two sets of circles commencing with C_1' and C_1'', the nearest circles to P, and let us denote by μ_i' the left-hand intersection of C_i' with L and by ν_i' the right-hand intersection, with the opposite conventions for μ_i'' and ν_i'', the intersections of C_i'' with L. The points $\mu_1', \cdots, \mu_{n'}', \mu_1'', \cdots, \mu_{n''}''$ may be obtained from the points $\alpha_1', \cdots, \alpha_{n'}', \alpha_1'', \cdots, \alpha_{n''}''$ by a similarity transformation with P as center, and as a line L' is allowed to rotate about P its intersections with C_i' and C_i'' have always this same property. In fact, we may consider properly chosen intersections of L with C_i' and C_i'' to be the roots of $\bar{f}_1$ and $\bar{f}_2$; then the roots of J trace the circles C_i. In particular, when L' is tangent to the circles $C_1', \cdots, C_{n'}', C_1'', \cdots, C_{n''}''$, the points of tangency have this relation to the points of tangency with $C_1, \cdots, C_n$.

To prove Theorem VI, we consider as usual the field of force given by Theorem I (I, p. 291 = I, p. 101). We can obtain immediately a qualitative idea of the locus of the roots of J. No point Q above both of the tangents T and T' common to the circles $C_1, \cdots, C_n, C_1', \cdots, C_{n'}', C_1'', \cdots, C_{n''}''$ can be a root of J. For the force at Q due to the particles situated at the roots of f_1 has a component normal to PQ and that due to the particles situated at the roots of f_2 has a component normal to PQ and in the same sense; at least one of these components is different from zero. Thus also no point below both T and T' can be a root of J and no point above T (or T') but not above T' (or T) yet lying to the right of C_1'' and to the left of C_1' can be a root of J. No point Q of T (or T') can be a root of J unless all the roots of f_1 and of f_2 are on T (or T') which can only occur if these roots lie at the points of tangency of T (or T') with the circles bounding their proper loci; that is, if Q lies at a point of tangency of T (or T') and a C_i. Inasmuch as the locus of the roots of J is a closed point set there must be some sort of a boundary of that locus between any two of the circles C_i.

By means of the similarity transformation with P as center, we see that every point of the locus as stated in Theorem VI is really a point of the locus. To complete the determination of the locus we have merely to prove that if a point Q is a point of the boundary of the locus, that point is on one of the circles C_i.

The interior and boundary of the circle C_i' or C_i'' which is the locus of more than one root of f_1 or of f_2 is also the interior and boundary of one of the circles C_i; every point interior to or on such a circle is a point of the locus of the roots of J. A point on or interior to two circles C_i' and C_{i+1}' (and so of course C_i'' and C_{i+1}'') is also on or within a circle C_j; in fact there is a root α_j of $\bar{J}$ between α_i' and α_{i+1}', and the circle C_j whose center is α_j contains in its interior the region common to C_i' and C_{i+1}', for this is true of any circle whose center lies between α_i' and α_{i+1}' if the circle has the two common tangents T and T'.

It merely remains to consider points Q exterior to all the circles $C_1', \cdots, C_{n'}', C_1'', \cdots, C_{n''}''$ or interior to or on at most one circle C_i' or C_i'' which is the locus of but one root of f_1 or f_2. In any case the particles at the roots of f_1 and f_2 may by Lemma I (II, p. 102) be considered to coincide in their respective loci so far as the force at Q is concerned.

9. **Proof of Theorem VI; replacing of two particles by a single particle.** The point Q is then to be considered as fixed, and for definiteness to lie to the right of P and of course above one of the lines T and T' but below the other, so that the line PQ actually cuts all the circles $C_1', \cdots, C_{n'}', C_1'', \cdots, C_{n''}''$. We know the loci of certain particles each representing several of the roots of f_1; we shall replace two of these particles by a single equivalent particle, then the new particle and a third of the original particles by a single equivalent particle, and so on until we have replaced all the particles representing the roots of f_1 by a single equivalent particle, and similarly for the particles representing the roots of f_2. A study of the properties of the loci of these various particles will enable us to prove that Q is on one of the circles C_i. We suppose for the present that Q is exterior to all the circles C_j'.

To replace two particles at ζ_1 and ζ_2 of masses m_1 and m_2 by a single particle ζ of mass $m_1 + m_2$ so that the force at Q shall be unchanged, we have the equation for ζ

$$\frac{m_1}{\zeta_1 - q} + \frac{m_2}{\zeta_2 - q} = \frac{m_1 + m_2}{\zeta - q},$$

where q is the complex number representing the point Q. This equation is equivalent to

$$\frac{(\zeta_1 - q)(\zeta_2 - \zeta)}{(\zeta_2 - q)(\zeta - \zeta_1)} = \frac{m_1}{m_2}.$$

We wish to replace the particles ζ_1 and ζ_2, whose loci are the interiors and boundaries of C_1' and C_2' and which represent all the roots of f_1 having these regions as loci, by a single equivalent particle. If we transform Q to infinity, we shall have precisely the conditions of III, Theorem VIII; m_1 and m_2 are both positive and the points ζ_1 and ζ_2 are always separated by ζ and Q. Then the variable circle C of Lemma IV (II, p. 105) moves so as always to cut C_1' and C_2' at the same angle, and cuts also S_1', the circle bounding the locus of ζ, also at this same angle. When Q is transformed back to the finite part of the plane, it remains true that C cuts C_1' and C_2' at the same angle; C cuts S_1' at this same angle or the supplementary angle according as the locus S_1' is interior or exterior to its bounding circle.* We leave this fact to be verified by the reader; this can be done by considering any one circle C under the

* This difference in behavior, which we shall constantly meet, disappears entirely if we project stereographically on to the sphere.

transformation of Q from the point at infinity to its original finite position. In particular it will be noticed that the line PQ is one of the circles C cutting C_1', C_2', S_1' at the proper angles. When ζ_1 and ζ_2 are on PQ and are chosen as the right-hand (left-hand) intersections of PQ with C_1' and C_2', ζ is on PQ and at the right-hand or left-hand (left-hand or right-hand) intersection of PQ with S_1' according as the locus S_1' is interior or exterior to its bounding circle. The converse of this statement is also true; such a choice of ζ leads to a unique determination of ζ_1 and ζ_2 as described.

We have supposed Q to be exterior to all the circles C_i'; suppose now Q exterior to C_1' but interior to C_2'; we need not consider Q interior to the two circles. When Q is transformed to infinity we have a special case of Theorem I, but no longer a special case of III, Theorem VIII. The circle C which generates as in Lemma IV (II, p. 105) the boundary of S_1' cuts C_1' and C_2' at supplementary angles. In fact, if we assume C to cut C_1' and C_2' at equal angles, but not at supplementary angles, when ζ_1, ζ_2, ζ are on the boundaries of their loci, the line $\zeta_1\zeta_2\zeta$ can be rotated about ζ so that ζ_1 and ζ_2 move into the interiors of their loci, so ζ cannot be on the boundary of its locus. The circle S_1' is found to be cut by C at an angle equal to that cut on C_2' and supplementary to that cut on C_1'.

We shall not consider in detail the case that Q is on C_1' or C_2'; we need not consider Q on both circles. Whether Q is on or within one circle or exterior to all the circles C_i' it is always true that when Q is in its original finite position C cuts C_1' and C_2' at the same angle, and cuts S_1' at this same angle or the supplementary angle according as the locus S_1' is interior or exterior to its bounding circle. When ζ_1 and ζ_2 are chosen properly as the intersections of C_1' and C_2' with PQ, one of the circles C, ζ is on PQ and on S_1', and conversely. The tangents to these three circles at those three points are parallel.

We have thus replaced the particles at ζ_1 and ζ_2 by a single equivalent particle. So far as the force at Q is concerned, we can replace ζ_1 and ζ_2 at any positions in their loci by ζ in its locus S_1', and for any position of ζ in S_1' we can determine ζ_1 and ζ_2 in their loci so that the force at Q is the same. If Q is in C_1' or C_2', the force at Q can be made as large as desired, so Q must be in S_1', and conversely. If Q is on C_1' or C_2', the force at Q can be made as large as desired but only in certain special directions, so Q is on S_1', and conversely. If the region S_1' is external to its bounding circle, the force at Q is zero for proper choice of ζ and hence of ζ_1 and ζ_2, and conversely.

Next we replace by a single equivalent particle ζ' the particle ζ as just determined and ζ_3, the particle which represents the roots of f_1 whose common locus is C_3' (assuming the existence of this set of roots). No further detailed discussion is required of this new situation; as before, a system of circles C' cuts S_1', C_3', and the boundary S_2' of the locus of ζ' at equal angles or at angles

supplementary to the angle cut on C_3' according as the loci S_1' and S_2' do not or do include the point at infinity. The line PQ is as before one of the system of circles C'. When ζ' is on PQ and on S_2', ζ_3 is on PQ and on C_3' and ζ is on PQ and S_1', so ζ_1 and ζ_2 are on PQ and C_1' and C_2' respectively; moreover, the tangents to C_1', C_2', C_3' at ζ_1, ζ_2, ζ_3, respectively, are parallel. The converse of this statement is also true.

We continue in this same manner to replace pairs of particles by equivalent particles, and finally replace all the particles $\zeta_1, \zeta_2, \cdots, \zeta_{n'}$ representing the roots of f_1 by a single particle η_1 whose locus is a circular region S_1 whose boundary is cut by PQ at an angle supplementary or equal to the angle cut on $C_1', C_2', \cdots, C_{n'}'$ according as the locus contains or does not contain the point at infinity. When η_1 is on PQ and on S_1, we know that $\zeta_1, \cdots, \zeta_{n'}$ are on $C_1', \cdots, C_{n'}'$ respectively and that the tangents to these circles at these points are parallel. It follows from reasoning to be given later that at no intermediate stage is the locus of one of our auxiliary particles the entire plane.

Similarly the particles $\xi_1, \xi_2, \cdots, \xi_{n''}$ representing the roots of f_2 are replaced by a single particle η_2 whose locus is a circular region S_2 which is either one of the points P or P' or is bounded by a circle S_2 which is cut by PQ at an angle equal to the angles cut on the circles $C_1'', C_2'', \cdots, C_{n''}''$. In fact, if all the roots of f_2 are not concentrated at P' the particles corresponding to the roots of f_2 always exert at Q a force not zero, so the locus of η_2 does not include the point at infinity. When η_2 is on PQ and S_2, we know that $\xi_1, \xi_2, \cdots, \xi_{n''}$ are on $C_1'', C_2'', \cdots, C_{n''}''$ respectively, and that the tangents to these circles at these points are parallel.

10. **Theorem VI: proof completed.** For Q to be a root of J, the loci S_1 and S_2 must have at least one point in common, at which point are to coincide η_1 and η_2 so that their resultant force at Q shall be zero. Such a common point cannot be Q, for Q is not a point of S_2. The loci S_1 and S_2 cannot overlap if Q is on the boundary of its locus, for when Q varies slightly in any direction, S_1 and S_2 vary but slightly. If S_1 and S_2 overlap, we may vary Q slightly in any direction but so little that S_1 and S_2 still have common points, so that Q remains a root of J for some choice of η_1 and η_2 and hence is not on the boundary of the locus of the roots of J. We defer until later the possibilities that the two circles S_1 and S_2 coincide or that S_1 or S_2 may be the entire plane or a point.

The regions S_1 and S_2 have but a single point in common, and since PQ cuts the circles S_1 and S_2 at equal or supplementary angles according as S_1 does or does not contain the point at infinity, a single point common to these two loci must lie on PQ. That is, η_1 and η_2 lie on S_1 and S_2 respectively, and on PQ, so $\xi_1, \xi_2, \cdots, \xi_{n'}, \xi_1, \xi_2, \cdots, \xi_{n''}$ all lie on PQ and the lines tangent

to the circles C_1', C_2', $\cdots$, $C_{n'}'$, C_1'', C_2'', $\cdots$, $C_{n''}''$ at these points are parallel. Then Q lies on one of the circles C_i.*

The possibility that the circles S_1 and S_2 coincide is readily treated. We may choose η_1 and η_2 to coincide on S_1 and on S_2, and on PQ. These points are still on the boundaries of their respective loci, and hence the previous reasoning is valid.

According to the assumptions already made, the locus S_1 cannot be a point. The locus S_2 will be a point when and only when the roots of f_2 are concentrated at P or P' or both. But in such a case the single point S_2 is on the line PQ and the preceding reasoning holds.

The possibility that S_1 or S_2 may be the entire plane remains to be considered. If one of these loci is the entire plane, the other must be a point; otherwise we have essentially the case of overlapping already disposed of. The locus S_2 is either the point P' or does not contain P', so is never the entire plane. If S_1 is the entire plane, we may suppose S_2 to be a point which of course lies on PQ. We prove our former result by a limiting process. When a point Q' is very near Q but external to the locus of the roots of J, the circles S_1', S_2', $\cdots$, S_1 are very near the corresponding circles for Q; for Q' the locus S_1 is certainly not the entire plane. Denote by Σ_2 the point at which is located the single particle representing all the roots of f_2, so far as concerns the force at Q'; Σ_2 is not in the locus S_1. When Q' approaches Q always remaining exterior to the locus of the roots of J, Σ_2 approaches the point S_2. The circle S_1 corresponding to Q' becomes smaller and smaller, the locus S_1 never contains Σ_2, so the circle S_1 approaches the point S_2. We may choose η_1 an intersection with PQ' of the circle S_1 corresponding to Q', and we shall have the points ζ_1, $\cdots$, $\zeta_{n'}$ on PQ' and on the circles C_1', $\cdots$, $C_{n'}'$. When Q' approaches Q, PQ' approaches PQ, the point η_1 approaches S_2 and the points ζ_1, $\cdots$, $\zeta_{n'}$ approach points on PQ and on C_1', $\cdots$, $C_{n'}'$. These limiting points can be taken as corresponding to η_1 coinciding with S_2 and thus give our result that Q lies on one of the circles C_1, $\cdots$, C_n.

Theorem VI is now completely proved except for its last sentence. When we notice the number of roots of J in a region C_i and remark that if the roots of f_1 and f_2 are varied continuously then the roots of J vary continuously, and that if C_i is exterior to every other circle C_j no root can enter or leave C_i, this last sentence is seen to be true. It hardly need be added that a number of circles C_i which may have common points but which have no point in common with any other circle C_j contain a number of roots of J equal to the sum of the multiplicities m_i corresponding to their centers as roots of J.

* The mere fact that the tangents at these points are parallel does not rule out certain isolated points Q on the line $P\alpha_1' \cdots \alpha_{n'}' \alpha_1'' \cdots \alpha_{n''}''$ but a more detailed consideration of the loci S_1', S_2', $\cdots$ does rule them out without difficulty.

11. **Generalization of Theorem VI by transformation.** In Theorem VI we have assumed that P is interior to none of the circles C_1', $\cdots$, $C_{n'}'$, C_1'', $\cdots$, $C_{n''}''$. The theorem is still true if this assumption is omitted, even if we permit roots of one or both forms to lie at infinity, except that the locus of the roots of J may be the entire plane, and will surely be the entire plane if P is interior to or on the circles C_1', $\cdots$, $C_{n'}'$, C_1'', $\cdots$, $C_{n''}''$, and if no roots of either ground form are constrained to lie at P'. The proof as given requires only a few minor modifications to apply to this new configuration. If the circles C_1', $\cdots$, $C_{n'}'$, C_1'', $\cdots$, $C_{n''}''$ all have the common center P, the circles bounding the locus of the roots of J are not determined by the position of their centers as described in the statement of Theorem VI, but are to be determined for example by their points of intersection with an arbitrary line through P, precisely as we considered the points μ_i' and μ_i'' in § 8.

Theorem VI has the advantage over Theorem II of being entirely symmetrical with respect to the two forms f_1 and f_2. The special case of Theorem VI where there are two circles C_1', C_2' and merely one circle C_1'' leads to merely one circle C_1 distinct from the three original circles. For this case, Theorems II and VI give the same result. But of course Theorem II is more general than this particular situation. Thus, the result for the jacobian problem of the theorem stated in II, pp. 114–115, or indeed of the problem of § 9 where Q is interior to C_2' is included in Theorem II but not in Theorem VI. There is, however, a general theorem concerned with an indefinite number of circular regions C_1', $\cdots$, $C_{n'}'$, C_1'', $\cdots$, $C_{n''}''$ which generalizes all possible situations of Theorem II and which is to be proved in §§ 13–15 by the methods we have been using.

Theorem VI as stated is not invariant under linear transformation. If we perform such a transformation we obtain the following new result:

THEOREM VII. *If the loci of the roots of f_1 are circular regions bounded by circles each of which is tangent internally to a circle L_1 and externally to a circle L_2 (tangent to both L_1 and L_2 internally) and if the loci of the roots of f_2 are circular regions bounded by circles each of which is tangent externally to L_1 and internally to L_2 (tangent to both L_1 and L_2 externally), and if these loci are so related to their bounding circles that they contain neither the entire circle L_1 nor the entire circle L_2, then the locus of the roots of the jacobian J of f_1 and f_2 is a number of circular regions bounded by circles each tangent internally to L_1 and externally to L_2 or tangent externally to L_1 and internally to L_2 (tangent to L_1 and L_2 internally or to L_1 and L_2 externally) and such that each region is so related to its bounding circle that it contains neither the entire circle L_1 nor the entire circle L_2. The exact location of the circles bounding the locus of the roots of J can be determined by allowing the roots of the ground forms to coincide on L_1 or on L_2, always remaining in their proper loci; the roots of J are the points of tangency with L_1*

and L_2 of the circles desired. The number of roots of J in any of these circular regions having no point in common with any other of these regions is the multiplicity as a root of J of the points of tangency of its boundary with L_1 and L_2 under these conditions.

There is a limiting case of this theorem yet different from Theorem VI where L_1 is a straight line and L_2 a proper circle, but we shall not state the result in detail.

12. **Application of Theorem VI to the roots of the derivatives of polynomials.** Theorem VI and in fact Theorem VII can be used to obtain results concerning the roots of the derivative of a rational function by means of the remark made in I, p. 297 (or II, p. 114). We thus consider $C_1', \cdots, C_{n'}'$ as the loci of the roots and $C_1'', \cdots, C_{n''}''$ as the loci of the poles of the given function. The regions $C_1, \cdots, C_n$ together with the possibility of the point at infinity appear as the loci of the roots of the derivative. There is a peculiar difference, however, between this locus and the corresponding locus of the roots of the jacobian. A region C_i'' which is the locus of more than one pole of the original function is the locus of at least one root of the derivative, with the exception that no point of the bounding circle C_i'' can be a root of the derivative unless it is interior to another region C_j''; we leave to the reader the verification of this statement.

We shall dwell at some length on an application of the above remark applied to Theorem VI, which is indeed a special case of that theorem, concerning the derivative of a polynomial:

THEOREM VIII. *Let the interiors and boundaries of circles $C_1, \cdots, C_n$ whose centers are $\alpha_1, \cdots, \alpha_n$ be the loci of $m_1, \cdots, m_n$ roots respectively of a polynomial $f(z)$ which has no other roots; suppose these circles to have a common external center of similitude P actually exterior to all these circles. Denote by $g(z)$ the polynomial $f(z)$ when all its roots are concentrated at the centers of their proper circles, and denote by $\alpha_1', \cdots, \alpha_{n'}'$ the distinct roots of its derivative $g'(z)$, of respective multiplicities $m_1', \cdots, m_{n'}'$. Then the locus of the roots of $f'(z)$ is composed of the interiors and boundaries of the circles $C_1', \cdots, C_{n'}'$ whose centers are $\alpha_1', \cdots, \alpha_{n'}'$ and whose radii are such that P is a common external center of similitude for the circles $C_1, \cdots, C_n, C_1', \cdots, C_{n'}'$. A circle C_i' which has no point in common with another circle C_j' contains m_i' roots of $f'(z)$.*

An extreme degenerate case of this theorem is when all the C_i are null circles, and $f(z)$ is identical with $g(z)$. The case of merely two circles brings us back to a theorem given in II, p. 115.

Inasmuch as the circles C_i' have P as a common external center of similitude, Theorem VIII can be applied again to the polynomial $f'(z)$ and shows that the roots of $f''(z)$ lie on or within certain circles $C_1'', \cdots, C_{n''}''$. The most obvious consideration of the geometric situation shows that any point on or within one of these circles actually is a point of the locus.

It is to be noticed, however, that this reasoning can be used only if we know that the circles C_i' contain respectively m_i' roots of $f'(z)$, and it is pointed out in § 4 that one of these circles does not necessarily contain precisely that number of roots of $f'(z)$ and in fact may contain no such root. Thus the reasoning can be used only if we know that the circles C_i' are mutually external. This is always the case for a given set of values of $\alpha_1, \cdots, \alpha_n$ if the circles C_i are sufficiently small, so in the following theorem we require that the circles C_i be sufficiently small. This means, more explicitly, that the theorem is true for a definite derivative $g^{(k)}$ if the circles C_i are so small that no circle $C_i^{(m)}$ has a point in common with a circle $C_j^{(m)}$ $(i \neq j)$, for $m = 1, 2, \cdots, k-1$.

Thus we have

THEOREM IX. *Let the interiors and boundaries of circles $C_1, \cdots, C_n$ whose centers are $\alpha_1, \cdots, \alpha_n$ be the loci of $m_1, \cdots, m_n$ roots respectively of a polynomial $f(z)$ which has no other roots; suppose these circles to have a common external center of similitude P actually exterior to all these circles. Denote by $g(z)$ the polynomial $f(z)$ when all its roots are concentrated at the centers of their proper circles, and denote by $\alpha_1^{(k)}, \cdots, \alpha_{n^{(k)}}^{(k)}$ the distinct roots of its kth derivative $g^{(k)}(z)$, of respective multiplicities $m_1^{(k)}, \cdots, m_{n^{(k)}}^{(k)}$. Then if the circles C_i are sufficiently small the locus of the roots of $f^{(k)}(z)$ is composed of the interiors and boundaries of the circles $C_1^{(k)}, \cdots, C_{n^{(k)}}^{(k)}$ whose centers are $\alpha_1^{(k)}, \cdots, \alpha_{n^{(k)}}^{(k)}$, and whose radii are such that P is a common external center of similitude for the circles $C_1, \cdots, C_n, C_1^{(k)}, \cdots, C_{n^{(k)}}^{(k)}$. A circle $C_i^{(k)}$ which has a point in common with no other circle $C_j^{(k)}$ contains precisely $m_i^{(k)}$ roots of $f^{(k)}(z)$.*

The special case of this theorem where there are but two of the original circles C_1 and C_2 has already been proved by another method.* For this special case we make no restriction on the size of the circles C_i.

A limiting case of Theorem IX is that P is infinite but the points $\alpha_1, \cdots, \alpha_n$ finite, and the radii of $C_1, \cdots, C_n$ finite. The circles $C_1, \cdots, C_n$ are then all equal. The theorem is true for this limiting case. In fact, suppose a root R of $f^{(k)}(z)$ to be exterior to all the circles $C_1^{(k)}, \cdots, C_{n^{(k)}}^{(k)}$. We can choose circles $S_1, \cdots, S_n$ having a finite point P as common external center of similitude and such that R is also exterior to all the circles $S_1^{(k)}, \cdots, S_{n^{(k)}}^{(k)}$ corresponding. This shows that every point of the locus is on or within $C_1^{(k)}, \cdots, C_{n^{(k)}}^{(k)}$; the converse is easily seen from translation of the situation for $g(z)$ and $g^{(k)}(z)$. This result may be expressed somewhat loosely as follows:

THEOREM X. *If the loci of the roots of a polynomial are the interiors and boundaries of sufficiently small equal circles whose centers lie on a line L, the locus of the roots of the kth derivative $f^{(k)}(z)$ consists of the interiors and bound-*

* Walsh, Paris Comptes Rendus, vol. 172 (1921), pp. 662–664.

aries of circles equal to these whose centers also lie on L and depend only on the centers of the original circles.

This new theorem for $k = 1$ is not a special case of Theorem VI and can easily be expressed in a form invariant under linear transformation, thus giving a new result for the jacobian of two binary forms (compare § 14) and for the derivative of a rational function.*

The approximate determination of the roots of the jacobian of two binary forms, of the derivative of a rational function, or of any derivative of a polynomial is thus made, by Theorems VI–X, to depend essentially on the determination of the roots of the jacobian, of the derivative of a rational function, or of any derivative of a polynomial which has all its roots real.

The extreme simplicity of Theorem X immediately raises the question of the truth of that theorem if the supposition of the collinearity of $\alpha_1, \cdots, \alpha_n$ is omitted. We can easily prove that the theorem is not true under this changed hypothesis by means of the remark of § 6. It is surely true under the changed hypothesis that every point on or within a circle C_i' that is equal to the C_i and whose center is α_i' is a point of the locus, but it is not true without further restrictions that the locus consists precisely of the points on and interior to the circles C_i'.

Consider a polynomial $g(z)$ with three simple roots $\alpha_1, \alpha_2, \alpha_3$, which are not collinear, so that neither root α_1', α_2' of $g'(z)$ is collinear with a pair of the points $\alpha_1, \alpha_2, \alpha_3$. Choose the equal circles C_1, C_2, C_3, with centers $\alpha_1, \alpha_2, \alpha_3$, of such small radius that for no possible choice of the points in their proper loci can we have two roots β_i and β_j of $f(z)$ collinear with a root β_i' of $f'(z)$. Suppose β_1' to be on C_1'; we choose β_i of such a nature that $\beta_i - \alpha_i = \beta_1' - \alpha_1'$. The circle C through $\beta_1, \beta_2, \beta_1'$ is not a straight line, the points β_3 and β_2 cannot satisfy the requirements of § 6 with regard to the circles C, C_1, and C_2, from which it follows that β_1' cannot be on the boundary of its locus.

13. **Theorem XI: an extension of Theorems II and VI.** We now come to the proof of the general theorem mentioned in § 11, which includes Theorem II as well as Theorem VI:

THEOREM XI. *Let f_1 and f_2 be two binary forms, and let circular regions $C_1', C_2', \cdots, C_{n'}'$; $C_1'', C_2'', \cdots, C_{n''}''$ be the respective loci of $m_1', m_2', \cdots, m_{n'}'$ roots of f_1 (which has no other roots) and of $m_1'', m_2'', \cdots, m_{n''}''$ roots of f_2 (which has no other roots). Suppose there is a family of coaxial circles S each of which cuts at the same angle all the circles C_i' which bound loci interior to them, and at*

* (Added in proof): It seems to the writer probable that Theorem IX is true with no restriction on the size of the circles C_i. The special case of this more general theorem where the circular regions C_i are half planes is for the case $k = 1$ contained by a limiting process in Theorem IX, and for all values of k has been established by Mr. B. Z. Linfield, in a paper to be published in these Transactions.

the supplementary angle all the circles C_i' which bound loci exterior to them, and which cuts at this same angle all the circles C_j'' which bound loci exterior to them and at the supplementary angle all circles C_j'' which bound loci interior to them. Then the locus of the roots of the jacobian of f_1 and f_2 is a number of circular regions bounded by circles $C_1, C_2, \cdots, C_n$ each of which is cut by every circle S at an angle equal or supplementary to the angles cut by S on C_i' and C_j''; the regions which are the loci of the roots of the jacobian may be either internal or external to their bounding circles. The circles $C_1, C_2, \cdots, C_n$ are included among the circles traced by the roots of the jacobian when all the roots of f_1 and f_2 are concentrated on the circles bounding their proper loci and move so that one of the circles S constantly passes through them all, while the lines tangent to these bounding circles C_i', C_j'' at the points which are the roots of f_1 and f_2 (and the lines tangent to C_k at the points which are the roots of the jacobian) all become parallel when S is transformed into a straight line. Any region C_i having no point in common with any other region C_j contains a number of roots of the jacobian equal to the multiplicity of the root of the jacobian which traces that circle C_i under these conditions.

We shall first undertake to prove this theorem for the simplest case, namely, that the circles S form a coaxial family having no point common to all those circles. We transform so that the circles S have a common center P. If the given circles C_i', C_j'' are all straight lines, all the regions C_i', C_j'' have a common point, the locus of the roots of the jacobian is the entire plane, and the theorem is proved. In any other case, all the circles $C_1', \cdots, C_{n'}', C_1'', \cdots, C_{n''}''$ are equal; the loci corresponding to the former can be considered to lie inside the bounding circles and those corresponding to the latter to lie outside the bounding circles.

We shall use the same method of proof as was used in § 9, namely, the replacing of all the particles in the field of force corresponding to the roots of f_1 by a single particle η_1, the replacing of all the particles corresponding to the roots of f_2 by a single particle η_2, and the study of the loci of η_1 and η_2. We shall prove that no matter what may be the location of the circular regions which are the loci or the distribution of the roots of the ground forms among these loci, *the locus of the roots of the jacobian is always the entire plane.*

If a point z is exterior to all the circles C_i', C_j'' and exterior to all the circles S which actually cut those circles C_i', C_j'', then z is a root of the jacobian. For if $n'' > 1$, z lies in C_1'' and C_2'', may lie at a multiple root of f_2, and hence is a point of the locus. If $n'' = 1$, replace the particles ζ_1 and ζ_2 of masses m_1' and m_2' whose loci are C_1' and C_2' by a single equivalent particle ζ of mass $m_1' + m_2'$ whose locus is a circular region S_1'. Then the circle S_1' is larger than C_1'; this follows from the fact that there are two circles through z tangent to S_1', C_1', C_2', and having the same kind of contact with all three of these circles; neither intersection of those two tangent circles with each other separates

any two of the points of tangency of the circles S_1', C_1', C_2'. The locus of ζ is the *interior* and boundary of S_1'. If $n' > 2$, we now replace ζ and the particle ζ_3 whose mass is m_3' and whose locus is C_3' by an equivalent particle ζ' of mass $m_1' + m_2' + m_3'$. The locus S_2' of ζ' is the interior of a circle which does not contain z and which is larger than C_3', by the reasoning just used. We continue in this way and finally replace all the particles representing the roots of f_1 by a single equivalent particle η_1 whose locus S_1 is larger than the circles C_i' (or if $n'' = 1$, equal to them). In any case, the region S_1 has at least one point in common with the region C_1'', so z is a point of the locus.

Denote by Σ_1 the larger and by Σ_2 the smaller of the two circles of the family S which are tangent to C_i' and C_j''. If z is interior to Σ_2, and if Σ_2 is exterior to C_i', C_j'', the reasoning just given applies with practically no change. If $n'' = 1$, the locus of η_1 which represents all the roots of f_1 is a region S_1 whose bounding circle is larger than the circles C_i', so the region S_1 must have at least one point in common with the region C_1'', and z is a point of the locus.

Let us now consider a point z between Σ_1 and Σ_2 under the assumption that Σ_2 is not interior to the circles C_i', C_j''. If $n'' = 1$, we find as before circles S_1', S_2', $\cdots$, S_1 all larger than C_1'. In fact, we need consider only points z interior to or on at most one circle C_i'. Describe a circle Σ through z and through the points of tangency of Σ_1 with C_1' and C_2'. When ζ_1 and ζ_2 lie at these two points, the point ζ corresponding lies on Σ, and is such that z and ζ separate ζ_1 and ζ_2. Hence ζ is exterior to Σ_1. Similarly there is a point ζ interior to Σ_2, so S_1' is indeed larger than C_1'. Thus we find z to be a point of the locus, for S_1 and C_1'' have a common point.

If $n'' > 1$, we need consider only points z interior to at most one circle C_i' and exterior to at most one circle C_j''. The circles S_1', S_2', $\cdots$, S_1 are all larger than C_1' (or equal to C_1' if $n' = 1$). The region S_1'' which is the locus of ξ, the particle equivalent to the particles ξ_1 and ξ_2 whose loci are C_1'' and C_2'', is a circular region which contains all the region common to C_1'' and C_2''. Hence the circle S_1'' is smaller than C_1''. The region S_1 which is the locus of the particle η_1 representing all the roots of f_1 is larger than each circular region C_i'. The region S_2 which is the locus of the particle η_2 representing all the roots of f_2 is bounded by a circle smaller than the bounding circle of each region C_i'', so S_1 and S_2 have at least one common point and z is a point of the locus.

It remains to consider points z interior to Σ_2, if Σ_2 is interior to C_i' and C_j'', but this treatment is so similar to the results already given that it is omitted. It remains also to consider points z on Σ_1 and on Σ_2, but since all other points of the plane are points of the locus and the locus of the roots of the jacobian is a closed point set, these points also belong to the locus. Theorem XI is now completely proved if the circles S have no point in common.

14. **Theorem XI, proof continued.** We next undertake to prove Theorem

XI for the case that the circles S form a coaxial family of circles all tangent at a single point P, which point we transform to infinity. If none of the original circular regions $C_1', \cdots, C_{n'}', C_1'', \cdots, C_{n''}''$ is the point at infinity, our results just proved for the case of circles S having no point in common hold without essential change; the entire plane is the locus of the roots of the jacobian. But P may be considered a null circle, the locus of a number of roots of f_1 and of f_2; in this case the entire plane need not be the locus of the roots of the jacobian.

If all the roots of f_2 are concentrated at P, then either all the roots of f_1 are also concentrated at P and the locus of the roots of the jacobian is the whole plane, or there are a number of fixed equal circles $C_1', \cdots, C_{n'}'$ bounding loci interior to them. In this latter case the field of force is precisely the field corresponding to Theorem X, so for this case Theorem XI is already proved. The case that there is at least one finite circle C_j'' requires some further consideration.

Denote by Σ_1 and Σ_2 the lines which belong to the coaxial family S and which are tangent to all the circles $C_1', \cdots, C_{n''}''$ and transform so that Σ_1 and Σ_2 are horizontal, with Σ_1 above Σ_2. Any point z on the boundary of the locus of the roots of the jacobian must lie on one of the circles traced by the roots of the jacobian when the roots of the ground forms trace the boundaries of their respective loci all constantly lying on one variable circle S and tracing the circles $C_1', \cdots, C_{n''}''$ in the same sense; this is the location of the roots of the ground forms described in the statement of Theorem XI. This fact is proved precisely as in §§ 9, 10, if z lies on a circle S which actually cuts the circles $C_1', \cdots, C_{n''}''$. We replace the particles at the roots of f_1 by a single equivalent particle η_1, the particles at the roots of f_2 by a single equivalent particle η_2, and notice that when z is on the boundary of its locus the loci of η_1 and η_2 cannot overlap. To complete the proof of Theorem XI in our special case it is sufficient to consider points z say above Σ_1 and to prove that all such points are points of the locus of the roots of the jacobian.

If z is a point above Σ_1 and if there are two or more finite circles C_j'', z is common to two or more of those circular regions and is therefore a point of the locus. If there is but one circle C_j'' other than at infinity, further consideration is required.

The locus of η_2 is the exterior of a circle S_2 obtained from C_1'' by similarity transformation with z as center, and S_2 is farther from z than is C_1''. Thus if there is but one finite circle C_1', the locus of η_1 is the interior of a circle S_1' obtained from C_1' by similarity transformation with z as center, and the loci S_1' and S_2 must have at least one common point, so z belongs to the locus.

If there are two finite circles C_1', we replace the particles whose loci are C_1' and C_2' by a single equivalent particle whose locus is S_1', and then replace that

particle and the particles at infinity which represent the roots of f_1 by a single equivalent particle whose locus is S_1. There are two circles through z tangent to S_1', C_1', C_2'; one of these circles contains S_1', C_1', C_2' and the other contains none of these circles. The external tangents to S_1' and C_1' intersect on the horizontal line through z, the radical axis of the circles through z tangent to S_1', C_1', C_2'.* Hence there is a circle Σ tangent to Σ_1 and Σ_2 and such that Σ and S_1' have z as common external center of similitude. It follows that the regions S_1 and S_2 have at least one common point. In fact, if there is no line through z which cuts both C_1'' and S_1', S_1 is entirely interior to S_2. If there is a line through z which cuts both C_1'' and S_1', a line through z and tangent to C_1'' cuts S_1' and S_1 and lies wholly in S_2. Thus z is a point of the locus.

If there are three finite circles C_i', we find S_1' as before; the external center of similitude of S_1' and C_3' lies on the horizontal line through z, so that line is the radical axis of the two circles through z and tangent to S_1', S_2', C_3'; one of these tangent circles contains S_1', S_2', C_3', and the other contains none of these three circles. Then the external center of similitude of S_2' and C_3' lies on the horizontal line through z, and as before we find that S_1 and S_2 have at least one common point. This reasoning is general for any number of circles C_i'.

Every point z above Σ_1 and hence every point below Σ_2 is a point of the locus. We may show either by similar reasoning or as in § 13 that every point on either of these lines belongs to the locus.

Theorem XI is thus proved for circles S all tangent at a single point. It is worth while, perhaps, to point out explicitly that there actually exist situations with one or more circles C_j'', and where the locus of the roots of the jacobian is not the entire plane. Thus, let there be merely two finite circles C_j'' and let z be interior to both of them. Then S_1'' is a region exterior to a circle which surrounds z. If z is exterior to all the circles C_i' and if the locus S_1 is interior to its bounding circle, it is possible so to choose the number of roots of f_2 at infinity that S_2 shall be the region exterior to a circle which entirely contains S_1. Then S_1 and S_2 have no common point and z is not a point of the locus of the roots of the jacobian.

15. **Theorem XI; completion of the proof.** The case that the circles S of Theorem XI form a coaxial family of circles through two distinct points P and P' remains to be dealt with. Transform P' to infinity. The points P and P' are considered as null circles and hence allowed to be loci of a number of roots of f_1 or f_2 or both. As in § 8 we may assume that there is at least one circle C_i' or C_j'' distinct from P and P'.

If the circles $C_1', \cdots, C_{n'}''$ surround P, Theorem XI can be proved precisely as in §§ 8–10. If $C_1', \cdots, C_{n'}''$ do not surround P, these same methods show that no point z is on the boundary of its locus unless z is on one of the circles

* Coolidge, *A Treatise on the Circle and the Sphere*, p. 111, Theorem 217.

described in the theorem, provided that there is a circle S through z which actually cuts all the circles C_1', $\cdots$, $C_{n''}''$. If all the finite regions C_1', $\cdots$, $C_{n''}''$ are interior to their bounding circles, the theorem is Theorem VI and completely proved. If two or more of these regions are exterior to their bounding circles, every point z not on a circle S which cuts all the circles C_1', $\cdots$, $C_{n''}''$ is a possible position of pseudo-equilibrium and hence a point of the locus. It remains to consider the case of such points z with merely one finite region, say C_1', exterior to its bounding circle. Let the line of centers of the circles C_1', $\cdots$, $C_{n''}''$ be horizontal and denote by Σ_1 and Σ_2 the common tangents to C_1', $\cdots$, $C_{n''}''$. Let C_1' lie to the left of P. We shall phrase the proof for $n' > 1$, $n'' > 1$.

The particles ζ_1 and ζ_2 whose loci are C_1' and C_2' are to be replaced by a particle ζ whose locus is a circular region S_1'. There are two circles through z tangent to C_1', C_2', S_1', one of which includes C_1' but not C_2', the other of which includes C_2' but not C_1'. If the locus S_1' is not the entire plane, it follows from a simple consideration of points ζ_1, ζ_2, ζ on the circle through z orthogonal to C_1' and C_2' that these two tangent circles include S_1' and exclude S_1' respectively. If the locus S_1' is the entire plane, the loci S_2', S_3', $\cdots$, S_1 are all the entire plane and z is a point of the locus.

These two tangent circles intersect on the line Pz,* and the circle S_1' lies to the left of Pz. The external center of similitude of C_1' and S_1' and the internal center of similitude of C_2' and S_1' lie on Pz. It is thus true that the external center of similitude of S_1' and any circle C_j'' lies on Pz and that the internal center of similitude of S_1' and any circle C_j' other than C_1' lies on Pz.

We now replace ζ and ζ_3, the particle whose locus is C_3', by a single equivalent particle ζ' whose locus is a circular region S_2'. If the locus S_2' is not the entire plane, there are two circles through z which intersect on Pz and which are tangent to S_1', C_3', S_2'; one of these tangent circles contains S_1' and S_2' but does not contain C_3', the other contains C_3' but neither S_1' nor S_2'. We continue in this way and finally reach a circle S_1 which bounds the locus of the point η_1 which represents all the roots of f_1; the locus of η_1 is exterior to S_1. The external center of similitude of S_1 and any of the finite circles C_j'' (and also of C_1') lies on Pz, and the internal center of similitude of S_1 and any of the finite circles C_j' except C_1' lies on Pz.

Similarly the particles representing the roots of f_2 are replaced by a single equivalent particle η_2 whose locus is the interior of a circle S_2 such that the external center of similitude of S_2 and any of the finite circles C_j' lies on Pz.

Hence the external center of similitude of S_1 and S_2 lies on Pz, from which it follows as before that there is at least one point common to S_1 and S_2, so z is a point of the locus. Likewise all points of Σ_1 and Σ_2 are points of the locus.

* Coolidge, loc. cit.

As in § 14, cases actually arise here where all the regions $C_1', \cdots, C_n''$ are not within their finite bounding circles and yet the locus of z is not the entire plane; the proof is as in § 14.

The number of roots of the jacobian in a region C_i which has no point in common with any other region C_i which is a part of the locus of the roots of the jacobian can be determined as in § 10 for Theorem VI; the proof of Theorem XI is now complete.

The determination in Theorem XI of whether or not a given circle C_i is actually a part of the boundary of the locus of the roots of the jacobian, and if so whether the circular region corresponding lies interior or exterior to C_i, can be made in any given case by the methods developed in the present chapter.

Theorem XI has obvious applications which will easily be made by the reader to the study of the location of the roots of the derivative of a polynomial and of the derivative of a rational function.

Chapter III: On centers of gravity

16. **The loci of certain centers of gravity.** There is a striking analogy between some of our results concerning the location of the roots of the jacobian of two binary forms and results which are easily proved concerning the location of the center of gravity of a number of particles. Thus, the fact that if a number of positive particles lie in a circle their center of gravity also lies in that circle is analogous to Lemma I (II, p. 102) and was used in the proof of that lemma, and is also analogous to the theorem of Lucas. From this fact and Theorem VIII of III we prove the analogue of a theorem given in II, p. 115 (= Theorem I of S) precisely as that theorem was proved:

Theorem XII. *If the interiors and boundaries of two circles C_1 and C_2 of centers α_1 and α_2 and radii r_1 and r_2 are the loci respectively of m_1 and m_2 unit positive particles, then the locus of the center of gravity of these particles is the interior and boundary of the circle C whose center is*

$$\frac{m_1 \alpha_1 + m_2 \alpha_2}{m_1 + m_2}$$

and whose radius is

$$\frac{m_1 r_1 + m_2 r_2}{m_1 + m_2}.$$

The three circles C_1, C_2, C have as common external center of similitude the point

$$\frac{r_1 \alpha_2 - r_2 \alpha_1}{r_1 - r_2}.$$

Theorem XII can be largely extended by the method of proof used for III, Theorem VIII in S:

THEOREM XIII. *If the interiors and boundaries of n circles C_i, whose centers are α_i and radii r_i, are the loci respectively of m_i unit positive particles, then the locus of the center of gravity of these particles is the interior and boundary of the circle C whose center is*

$$\alpha = \frac{m_1\alpha_1 + m_2\alpha_2 + \cdots + m_n\alpha_n}{m_1 + m_2 + \cdots + m_n}$$

and whose radius is

$$r = \frac{m_1 r_1 + m_2 r_2 + \cdots + m_n r_n}{m_1 + m_2 + \cdots + m_n}.$$

Denote the m_i particles in or on C_i by $z_1^{(i)}, z_2^{(i)}, \cdots, z_{m_i}^{(i)}$, so that $|z_j^{(i)} - \alpha_i| \leqq r_i$ for every i and j. The center of gravity of all the $m_1 + m_2 + \cdots + m_n$ particles is

$$z = \frac{(z_1^{(1)} + z_2^{(1)} + \cdots + z_{m_1}^{(1)}) + \cdots + (z_1^{(n)} + z_2^{(n)} + \cdots + z_{m_n}^{(n)})}{m_1 + m_2 + \cdots + m_n},$$

so that we have

$$z - \alpha = \frac{\begin{array}{l}[(z_1^{(1)} - \alpha_1) + (z_2^{(1)} - \alpha_1) + \cdots + (z_{m_1}^{(1)} - \alpha_1)] + \cdots \\ \quad + [(z_1^{(n)} - \alpha_n) + (z_2^{(n)} - \alpha_n) + \cdots + (z_{m_n}^{(n)} - \alpha_n)]\end{array}}{m_1 + m_2 + \cdots + m_n}$$

and hence z is on or within C.

Conversely, if z is given on or within C, we determine $z_j^{(i)}$ by the relation

$$z_j^{(i)} - \alpha_i = (z - \alpha)\frac{r_i}{r},$$

and we have the $z_j^{(i)}$ satisfying the proper conditions. The proof is thus complete. It may be remarked that when the $z_j^{(i)}$ trace their proper circles in such a manner that $(z_j^{(i)} - \alpha_i)/r_i$ is the same for every i and j, then this common value is equal to $(z - \alpha)/r$ while z traces its circle C.

Theorem XIII can be extended without difficulty in various directions: to particles of negative or even complex mass; to space of any number of dimensions; to give a result which shall be invariant under linear transformation; to regions other than the interiors of circles, especially convex regions. In this last extension, use is made of the fact that if m_i particles lie in a convex region their center of gravity also lies in that region; hence such results as III, Theorem IX can be applied.

There is much more than a mere analogy between Theorems XII and XIII for centers of gravity and our previous results concerning the derivatives of polynomials. In fact, the only root of the $(n-1)$st derivative of a polynomial of degree n lies at the center of gravity of the roots of that polynomial. When viewed in this light, Theorems XII and XIII are results relating to the

location of the roots of the derivatives of a polynomial even if not of the jacobian of two binary forms, and are conceived in precisely the same spirit as is Theorem IX. Thus the entire discussion of § 5 holds practically without change if we consider the problem of determining the locus of the roots of the kth derivative of a polynomial of degree n whose roots have certain assigned circular regions as their loci. Theorem XIII gives the complete solution of that problem for $k = n - 1$ if the assigned circular regions are interior to their bounding circles.

As a particular case of Theorem XIII, the fact that if a number of particles lie in a convex region their center of gravity also lies in that region follows from the theorem of Lucas* as applied successively to the various derivatives of a polynomial.

17. **The center of gravity of the roots of the derivative of a rational function and of the jacobian of two binary forms.** The center of gravity of any set of points has interesting properties with reference to that point set. It furnishes, for example, an approximate idea of the location of those points. Any line through the center of gravity either passes through all the points of the set or separates at least two of them.† We shall now find some results connecting the centers of gravity of related polynomials of the sort we have been considering. A classical theorem of this nature follows from a remark previously made:

THEOREM XIV. *The center of gravity of the roots of a polynomial coincides with the center of gravity of the roots of the derived polynomial.*

We derive the corresponding result for a rational function, which we take in the form

$$f(x) = \frac{x^m + a_0 x^{m-1} + a_1 x^{m-2} + \cdots + a_{m-1}}{x^n + b_0 x^{n-1} + b_1 x^{n-2} + \cdots + b_{n-1}};$$

$$f'(x) = \frac{(x^n + b_0 x^{n-1} + \cdots)(mx^{m-1} + (m-1)a_0 x^{m-2} + \cdots) - (x^m + a_0 x^{m-1} + \cdots)(nx^{n-1} + (n-1)b_0 x^{n-2} + \cdots)}{(x^n + b_0 x^{n-1} + \cdots)}.$$

If we denote by α the center of gravity of the finite roots of $f(x)$ and by β that of the finite poles of $f(x)$, if $m \neq n$ and if $f(x)$ has no finite multiple poles, we have for the center of gravity of the finite roots of $f'(x)$ the formula

$$\gamma = -\frac{(m-n-1)a_0 + (m-n+1)b_0}{(m+n-1)(m-n)}$$

$$= \frac{m(m-n-1)\alpha + n(m-n+1)\beta}{(m+n-1)(m-n)},$$

* On the other hand, Lemma I (II, p. 102) enables us to use this fact to give immediately a very simple proof of the theorem of Lucas.

† An application of this fact to the more precise location of the roots of algebraic equations is given by Laguerre, *Œuvres*, vol. I, pp. 56, 133.

which is a point collinear with α and β. If α and β coincide, γ coincides with them; if $m = n + 1$, γ coincides with β; if $m = n - 1$, γ coincides with α. If $n = 0$, we have Theorem XIV. If $m = n$, γ cannot be expressed in terms of merely α and β, as is found simply by computing γ.

Let us inquire in what respect this work on centers of gravity can be made invariant under linear transformation and can be applied to the jacobian of two binary forms.

The concept *center of gravity* is surely not invariant under linear transformation. In fact, given any two distinct points of the plane ξ and η, any third point ζ of the plane can, by a suitable transformation, be made to correspond to the center of gravity of the transformed ξ and η. We need simply to transform to infinity the harmonic conjugate of ζ with respect to ξ and η.

We cannot expect to obtain results with the ordinary definition of center of gravity, so we introduce a new definition. The point G is said to be the *centroid of a set of points with respect to* P if when P is transformed to infinity G transformed into the center of gravity of the points corresponding to the original set. We suppose that P is not a point of that set. It should be noted by way of justification of the definition that the point G is uniquely defined, since the center of gravity is invariant under similarity transformation. The relation between the points P and G is not reciprocal.

The centroid with respect to a point of a set of points gives a rough indication of the distribution of that set of points, like the ordinary center of gravity. In particular, if P is external to a circular region containing the set of points, G is also in that circular region. In fact, examination of the proof of Lemma I (II, p. 102) will show that the force at a point P external to a circular region C due to k particles in C is equivalent to the force at P due to k particles which coincide at a point Q in C, and Q is the centroid of the k particles with respect to P. Thus we are studying the relation between P, Q, and the k particles, which is the same as Laguerre's relation set up between those points, referred to in § 2.

Let f_1 and f_2 be two binary forms, of respective degrees p_1 and p_2, and let the point P at infinity be a k-fold root of f_1. Let α, β, γ be the centroids with respect to P of the roots of f_1 other than P, the roots of f_2 (all of which are supposed finite), the finite roots of the jacobian of f_1 and f_2, respectively. We easily find that

$$\gamma = \frac{(p_1 - k)(k+1)\alpha - (p_1 - kp_2)\beta}{k(p_1 + p_2 - k - 1)},$$

a point collinear with α and β. If α and β coincide, γ coincides with them; if $p_1 = kp_2$, $\gamma = \alpha$; if $p_1 = k$, $\gamma = \beta$, which is Theorem XIV. Always we shall have*

* It might seem at first sight that this cross ratio should be p_1 / k, since by Lemma II

$$(P, \alpha, \beta, \gamma) = \frac{(p_1 - k)(k+1)}{k(p_1 + p_2 - k - 1)},$$

which expresses the entire result in invariant form.

Chapter IV: On the roots of the jacobian of two real forms

18. **The locus of the roots of the derivatives of a polynomial whose roots are real.** The present chapter is devoted mainly to general theorems of the kind developed in Chapter II, but where we restrict ourselves to ground forms whose coefficients are real or can be made real by suitable linear transformation. We are placing additional restrictions on our ground forms, so it is to be expected that some additional properties will appear.

Any result concerning the location of the roots of the derivative of a polynomial is also a result concerning the roots of the jacobian of two binary forms. Thus all the facts proved in A can be given this interpretation and other results can be found by linear transformation.* The reader can easily formulate these new theorems. We now prove a new result concerning the derivatives of polynomials all of whose roots are real.

Theorem XV. *Let intervals I_i $(i = 1, 2, \cdots, m)$ of the axis of reals, whose end points are α_i, β_i, $\alpha_i \leqq \beta_i$, be the respective loci of m_i roots of a polynomial $f(z)$ which has no other roots. Then the locus of the roots of $f^{(k)}(z)$ is composed of a number of intervals $I_i^{(k)}$ of the axis of the reals. The left-hand end points of the intervals $I_i^{(k)}$ are the roots of $f^{(k)}(z)$ when the roots of $f(z)$ are concentrated at the points α_i; the right-hand end points are the corresponding roots of $f^{(k)}(z)$ when the roots of $f(z)$ are concentrated at the points β_i. Any interval $I_i^{(k)}$ which has no point in common with any other interval $I_j^{(k)}$ contains a number of roots of $f^{(k)}(z)$ equal to the multiplicity of its left-hand end point as a root of $f^{(k)}(z)$ when the roots of $f(z)$ are the points α_i. If the intervals I_i are all of equal length, the intervals $I_i^{(k)}$ are of this same length. If there is a point P which is a center of similitude for every pair of the intervals I_i (which is always true if $m = 2$), P is also a center of similitude for every pair of intervals $I_i^{(k)}, I_j^{(k)}$.*†

We prove this theorem under the assumption that no interval I_i reduces to

(II, p. 102) when the $p_1 - k$ finite roots of f_1 coalesce at α and the p_2 finite roots of f_2 coalesce at β there is but one position of equilibrium, namely, at the point γ' such that $(P, \alpha, \beta, \gamma') = p_1/k$. However, the jacobian vanishes not only at γ' but also at α and β if $p_1 - k$ and p_2 are greater than unity. It is the centroid with respect to P of *all* the finite roots of the jacobian that we have denoted by γ. The two formulas are the same when $p_1 - k = 1$, $p_2 = 1$.

*(Added in proof): There is an error in the statement of the italicized theorem of A, p. 133, as has been pointed out by Nagy, Jahresberichtder Deutschen Mathematiker-Vereinigung. vol. 31 (1922), pp. 245, 246. That theorem has no meaning as it appears at present, but becomes correct if the word *exterior* is replaced by the word *other*. The theorem is correctly stated in the abstract of A, Bulletin of the American Mathematical Society, vol. 26 (1919–20), p. 259.

†Some special cases of this theorem are given by Nagy, Jahresbericht der Deutschen Mathematiker-Vereinigung, vol. 27 (1918), pp. 37–43; 44–48. The special case $m = 2, k = 1$ is Theorem II of S.

a point; to include this more general case requires merely a slight change in phraseology. We prove the theorem first for the case $k = 1$. In the theorem the intervals are assumed to be finite, but the theorem can be extended to include infinite intervals.

Let us denote by $\alpha_i^{(k)}$ the roots of $f^{(k)}(z)$ when the roots of $f(z)$ are concentrated at the points α_i, $\alpha_i^{(k)} \leqq \alpha_j^{(k)}$ when $i < j$, and similarly by $\beta_i^{(k)}$ the roots of $f^{(k)}(z)$ when the roots of $f(z)$ are concentrated at the points β_i, $\beta_i^{(k)} \leqq \beta_j^{(k)}$ when $i < j$. The intervals $I_i^{(k)} : (\alpha_i^{(k)}, \beta_i^{(k)})$ are then to be proved to form the locus of the roots of $f^{(k)}(z)$.

Let us start with the roots of $f(z)$ concentrated at the points α_i, and move these roots continuously to the right until they reach the points β_i. The roots of $f'(z)$ also vary continuously in their totality; they start at the points α_i' and reach the points β_i'. If we number these roots, commencing at the left, we can even say that the nth root z_n' of $f'(z)$ varies continuously. We now prove that z_n' moves always to the right.

The equation determining z_n' is of the form

$$F = \frac{m_1}{z_n' - \gamma_1} + \frac{m_2}{z_n' - \gamma_2} + \cdots + \frac{m_m}{z_n' - \gamma_m} = 0, \tag{3}$$

where the γ_i are the roots of $f(z)$, coinciding in any multiplicities m_i desired. We compute the values

$$\frac{\partial F}{\partial z_n'} = -\frac{m_1}{(z_n' - \gamma_1)^2} - \frac{m_2}{(z_n' - \gamma_2)^2} - \cdots - \frac{m_m}{(z_n' - \gamma_m)^2},$$

$$\frac{\partial F}{\partial \gamma_i} = \frac{m_i}{(z_n' - \gamma_i)^2}.$$

It is always true that $\partial z_n'/\partial \gamma_i$ is positive, so z_n' always increases with γ_i.

Equation (3) is no longer valid to determine z_n' if z_n' is located at a multiple root of $f(z)$. Under these circumstances, if γ_i does not coincide with z_n', the motion of γ_i does not change the position of z_n'. If γ_i does coincide with z_n' and if γ_i is moved to the right, z_n' is either unchanged or moved to the right; this follows immediately from the fact that a k-fold root of $f(z)$ is a $(k-1)$-fold root of $f'(z)$ and from the fact that every interval bounded by roots of $f(z)$ contains at least one root of $f'(z)$.

From the general fact, then, that the nth root z_n' of $f'(z)$ varies continuously and in one sense under the indicated variation in the roots of $f(z)$, it follows that z_n' traces the entire interval from α_n' to β_n'. It remains to be shown that z_n' can never be outside of the interval (α_n', β_n'). If we assume z_n' to lie outside of that interval, say for definiteness to the right, for some possible position of the roots of $f(z)$, motion of those roots of $f(z)$ to the right always within their proper loci would move z_n' to the right and when the roots of $f(z)$ reached

the ends of their proper intervals z_n' would lie to the right of β_n', which is impossible.

The determination of the locus in Theorem XV is now complete for $k = 1$; the statement relative to the number of roots of $f'(z)$ in the various intervals is readily proved by the continuity methods previously used.

For the case of $k = 2$, the continuity of the motion of the roots of $f''(z)$ due to the motion of the roots of $f(z)$ shows that every point of each of the intervals I_k'' is a root of $f''(z)$ for some $f(z)$. No other point can be a root of $f''(z)$, for when the roots of $f(z)$ vary continuously in one sense, the roots of $f'(z)$ and therefore of $f''(z)$ vary continuously in that same sense. The number of roots of $f''(z)$ proper to the various intervals is as indicated. Continuance of the method of reasoning enables us to determine the locus for $k = 3$ and so on for the other values of k.

If all the intervals I_i are of the same length, the $f(z)$ whose roots are the β_i is obtained from the $f(z)$ whose roots are the α_i by a translation, so the β_i are obtained from the corresponding α_i by the same translation and the I_i' (and hence the $I_i^{(k)}$) are all of the same length as the I_i. If the β_i are obtained from the α_i by a similarity transformation, the $\beta_i^{(k)}$ are obtained from the $\alpha_i^{(k)}$ by the same transformation.

19. **The extension of theorems for the derivative of a polynomial to the roots of the jacobian.** Theorem XV cannot be immediately extended to the location of the roots of the jacobian of two binary forms, where the loci of the roots of both forms are intervals of the axis of reals. First, all the roots of both forms may coincide, so that the locus of the roots of the jacobian is not a number of intervals of the axis of reals. Second, the jacobian may have non-real roots even when it does not vanish identically.*

We can avoid this first possibility by requiring that the loci of the roots of f_1 and f_2 be so arranged that the two forms cannot be identically equal. We can avoid the second possibility by requiring that these loci be so arranged that no two roots of f_1 can separate two roots of f_2. Then all the roots of the jacobian are real, for on any interval bounded by roots of either form and containing no root of the other form there lies at least one root of the jacobian.

With these new restrictions, Theorem XV extends directly to the jacobian of two binary forms. If all the intervals which are the loci of the roots of both forms are finite, we consider the α_i (β_i) to be at the left-hand (right-hand) ends of those intervals which are loci of the roots of f_1 and at the right-hand (left-hand) ends of those intervals which are loci of the roots of f_2. For infinite intervals this notation is reversed. The locus of the roots of the jacobian is composed of the intervals whose end points are the corresponding

* This is shown by the simplest examples, such as $f_1 = z_1^2 - z_2^2$, $f_2 = z_1 z_2$, $J = 2(z_1^2 + z_2^2)$.

roots of the jacobian when the roots of the ground forms are respectively the α_i and the β_i.

A special case of this result is so similar to Theorem II that it deserves to be stated explicitly:

THEOREM XVI. *Let f_1 and f_2 be binary forms of degrees p_1 and p_2 respectively, and let arcs A_1, A_2, A_3 of a circle C be the respective loci of m roots of f_1, the remaining $p_1 - m$ roots of f_1, and all the roots of f_2. Suppose A_3 to have not more than one point in common with A_1 nor with A_2 and no point in common with both A_1 and A_2. Denote by A_4 the arc of C which is the locus of points z_4 such that*

$$(z_1, z_2, z_3, z_4) = \frac{p_1}{m},$$

when z_1, z_2, z_3 have the respective loci A_1, A_2, A_3. Then the locus of the roots of the jacobian of f_1 and f_2 is composed of the arc A_4 together with the arcs A_1, A_2, A_3, except that among the latter the corresponding arc is to be omitted if any of the numbers m, $p_1 - m$, p_2 is unity. If an arc A_i $(i = 1, 2, 3, 4)$ has no point in common with any other of those arcs which is a part of the locus of the jacobian, it contains precisely $m - 1$, $p_1 - m - 1$, $p_2 - 1$, or 1 of those roots according as $i = 1, 2, 3,$ or 4.

Theorem XVI, as a special case of our more general result on the location of the roots of the jacobian, needs no separate proof, but it is interesting to notice that it can be proved in precisely the same manner as Theorem II was proved. Theorem I in the proof of Theorem II is replaced by III, Theorem IV, and Lemma I (II, p. 102) is replaced by the following

LEMMA. *The force at a point P on a circle C due to k unit positive particles lying on an arc A of C not containing P is equivalent to the force at P due to k coincident particles lying on A.*

We shall now obtain a result which has some relation to Theorem XVI as well as to Jensen's theorem, proved in A. We are dealing with pairs of points inverse with respect to a line, and as in A shall term circles whose diameters are the segments joining such pairs of points *Jensen circles*. Let f_1 and f_2 be two real forms which have not necessarily all their roots real. Let finite or infinite segments I_1, I_2, I_3 of the axis of reals either contain respectively m roots of f_1, the remaining $p_1 - m$ roots of f_1, and the p_2 roots of f_2, or contain some of these roots and the intercepts on the axis of reals of the Jensen circles of the remainder. Then any real root of the jacobian of f_1 and f_2 which is exterior to I_1, I_2, and I_3 lies in the interval I_4 which is the locus of the point z_4 defined by

$$(z_1, z_2, z_3, z_4) = \frac{p_1}{m}$$

when I_1, I_2, I_3 are the respective loci of z_1, z_2, z_3.

To prove this result we require the preceding lemma and the fact that the force at a point due to two particles is equivalent to the force at that point due to two coincident particles situated at the harmonic conjugate of that point with respect to the other two. If a point P is exterior to I_i, its harmonic conjugate with respect to two points the intersections of whose Jensen circle with the axis of reals lie in I_i, also lies in I_i.

This result is not expressed so as to be invariant under linear transformation, for if a real linear transformation is made and if the point at infinity is not invariant the Jensen circles are not invariant.

Our final result on the roots of the jacobian is similarly not invariant under linear transformation; it can be proved from the fact proved in A, § 2, that the force due to two positive particles at a point above the axis of reals but interior to their Jensen circle has a component vertically downward; at a point above the axis of reals but exterior to the Jensen circle the force has a component vertically upward.

THEOREM XVII. *If the forms f_1 and f_2 are both real and if f_1 has no finite real root, there is no root of the jacobian of f_1 and f_2 exterior to all the Jensen circles corresponding to the roots of f_2 but interior to all the Jensen circles corresponding to the roots of f_1.*

20. **Conclusion: extension of results to other types of polynomials.** We have considered in this paper generalizations of Theorem II in various directions. There is still another direction which we have not mentioned, namely, to the roots of polynomials other than the jacobian of two binary forms or the derivatives of a polynomial.

Thus the jacobian of two forms f_1 and f_2, of respective degrees p_1 and p_2, all of whose roots are finite and which correspond to two polynomials ϕ_1 and ϕ_2, has the same roots as the polynomial

$$p_2 \phi_1' \phi_2 - p_1 \phi_1 \phi_2'.$$

If we set ϕ_1 equal to the product of two polynomials ψ_1 and ψ_2 of respective degrees m and $p_1 - m$, Theorem II refers to the roots of the polynomial

$$p_2 \psi_1' \psi_2 \phi_2 + p_2 \psi_1 \psi_2' \phi_2 - p_1 \psi_1 \psi_2 \phi_2', \tag{4}$$

when the roots of ψ_1, ψ_2, ϕ_2 have the respective loci C_1, C_2, C_3.

We shall generalize Theorem II by considering three polynomials ω_1, ω_2, ω_3, of respective degrees μ_1, μ_2, μ_3, whose roots have the respective loci C_1, C_2, C_3. Our conclusion concerns the polynomial

$$\lambda_1 \omega_1' \omega_2 \omega_3 + \lambda_2 \omega_1 \omega_2' \omega_3 + \lambda_3 \omega_1 \omega_2 \omega_3', \tag{5}$$

where λ_1, λ_2, λ_3 are real* numbers not all zero such that

$$\lambda_1 \mu_1 + \lambda_2 \mu_2 + \lambda_3 \mu_3 = 0.$$

* Proof of Theorem I for complex λ enables us to remove this restriction of reality. See Walsh, Rendiconti del Circolo Matematico di Palermo, vol. 46 (1922), pp. 236–248.

It will be noticed that the polynomial (5) is indeed a generalization of (4), and has the additional advantage of being symmetric in $\omega_1, \omega_2, \omega_3$.

If a point z is a root of (5) yet exterior to C_1, C_2, C_3, we must have

$$\lambda_1 \frac{\omega_1'}{\omega_1} + \lambda_2 \frac{\omega_2'}{\omega_2} + \lambda_3 \frac{\omega_3'}{\omega_3} = 0,$$

$$\frac{\lambda_1 \mu_1}{z - \alpha_1} + \frac{\lambda_2 \mu_2}{z - \alpha_2} + \frac{\lambda_3 \mu_3}{z - \alpha_3} = 0,$$

by Lemma I (II, p. 102), where $\alpha_1, \alpha_2, \alpha_3$ lie in C_1, C_2, C_3 respectively. Hence z is given by the cross ratio

$$(\alpha_1, \alpha_2, \alpha_3, z) = -\frac{\lambda_3 \mu_3}{\lambda_1 \mu_1},$$

and lies in the region C_4 of Theorem I corresponding to the value

$$\lambda = -\frac{\lambda_3 \mu_3}{\lambda_1 \mu_1}.$$

We leave it to the reader to verify that the locus of the roots of (5) is composed of C_4 together with the regions C_1, C_2, C_3, except that among the latter the corresponding region is to be omitted if any of the degrees μ_1, μ_2, μ_3 is unity. If a region C_i ($i = 1, 2, 3, 4$) has no point in common with any other of those regions which is a part of the locus of the roots of the jacobian, it contains precisely $\mu_1 - 1, \mu_2 - 1, \mu_3 - 1, 1$ of those roots according as $i = 1, 2, 3, 4$.

Many of the other theorems of the present paper, such as Theorems VI–XI, can similarly be extended to polynomials other than the jacobian of two binary forms or the derivative of a polynomial. Modifications of the methods used here can be made to apply to a still much broader type of polynomial about which the writer hopes to give some further results.

Catonsville, Md.

I-CONJUGATE OPERATORS OF AN ABELIAN GROUP*

BY

G. A. MILLER

I. Introduction

Two operators of any group G are said to be I-conjugate if they correspond in at least one of the possible automorphisms of G. Every characteristic subgroup of G includes all the I-conjugates of each of its operators, and if a subgroup includes all the I-conjugates of its operators it is characteristic. In the present article it will be assumed that G is abelian. As two operators of any abelian group are I-conjugate if their prime power constituents have this property, and vice versa, it will only be necessary to consider the case when the order of G is of the form p^m, p being a prime number. Hence this will be done in what follows unless the contrary is stated.

Two fundamental questions in regard to the I-conjugate operators of G are: how many sets of I-conjugate operators are there in any abelian group, and how many operators are found in each of these sets? Both of these questions are answered in what follows, and the method for determining these numbers which is developed here seems to be as direct as possible. It is evident that every two I-conjugate operators are of the same order, and that a necessary and sufficient condition that every two operators of G which are of the same order be also I-conjugate is that all the invariants of G be equal to each other.

Two definitions of independent generators of G are in common use. According to one of these definitions the operators $s_1, s_2, \cdots, s_\lambda$ are called a set of independent generators of G whenever they satisfy the two conditions that they generate G and that no $\lambda - 1$ of them generate G. The number λ is known to be an invariant of G. According to the second definition, these λ operators must satisfy the additional condition that the subgroup generated by an arbitrary subset of them have only identity in common with the subgroup generated by the rest of these operators. According to the first definition, all the operators of G which do not appear in any one of the possible sets of independent generators of G constitute a subgroup of G known as its ϕ-subgroup, while these operators constitute such a subgroup according to the second definition when and only when the ratio of the largest invariant to the smallest invariant of G does not exceed p.

* Presented to the Society, December 30, 1920.

To distinguish between sets of independent generators satisfying the first, or also the second, of these definitions, the latter are called *reduced sets of independent generators*. The former sets of independent generators are usually the most convenient when only questions relating to subgroups are considered, while the latter are more convenient in the study of conjugacy. In the present article it will be assumed that the sets of independent generators under consideration are reduced sets unless the contrary is stated. It will be seen that all the operators of G which do not appear in any such set generate a subgroup which includes all the independent generators of G except those of highest order and those whose order is equal to this highest order divided by p, if any of the latter exist.

When G has independent generators of different orders, its independent generators which are of the same order are evidently I-conjugate and can be selected from a set of I-conjugate operators of G which has no operator in common with the group generated by the remaining independent generators of the set. The latter independent generators can usually be selected in a large number of different ways and the subgroups which such operators generate may differ, but none of the subgroups can involve an operator of the set of I-conjugate operators from which the former independent generators must be chosen.

The number of the subgroups of G which are separately generated by all the independent generators of G which are of the same fixed order in its various possible sets of independent generators can easily be determined. In fact, it is the quotient obtained by dividing the number of ways in which the independent generators of such a subgroup can be selected from the operators of G by the number of ways in which these generators can be selected from the operators of one of these subgroups. If p^α is the order of such an independent generator the totality of the operators of order p^β, $0 \leq \beta \leq \alpha$, contained in all of these subgroups constitutes a single set of I-conjugate operators of G. Hence the distinct operators in all of these subgroups constitute the operators of α I-conjugate sets of G excluding identity. Two such subgroups corresponding to different values of α can have only identity in common, and G is the direct product of an arbitrary set of subgroups such that one and only one of the subgroups of this set corresponds to a particular possible value of α.

II. I-REDUCED OPERATORS OF A GROUP

It has been noted that when G contains a set of independent generators composed of λ_1 operators of order p^{α_1}, λ_2 operators of order p^{α_2}, $\cdots$, λ_γ operators of order p^{α_γ}, so that

$$\lambda_1 \alpha_1 + \lambda_2 \alpha_2 + \cdots + \lambda_\gamma \alpha_\gamma = m,$$

then the number of its sets of I-conjugate operators which are separately

powers of possible independent generators is $\alpha_1 + \alpha_2 + \cdots + \alpha_\tau$, exclusive of identity. Each of these sets contains one and only one operator which satisfies both of the following conditions: It is the lowest possible power of an operator in the set of independent generators $s_1, s_2, \cdots, s_\lambda$ which appears in the former set, and this operator has the smallest possible subscripts. Such an operator will be called an *I-reduced operator* and hence each of the given sets of I-conjugate operators contains one and only one I-reduced operator, and this is a power of an independent generator of G.

In general, an I-reduced operator is defined as the single operator of a set of I-conjugate operators of G which satisfies the following conditions: It involves powers of the smallest possible number of constituents which are separately powers of the operators $s_1, s_2, \cdots, s_\lambda$ for the set of I-conjugate operators in which it is found, each of these constituents is raised to the lowest possible power, and the subscripts of operators of the set $s_1, s_2, \cdots s_\lambda$ of which these constituents are powers are as small as possible. A necessary and sufficient condition that an I-reduced operator involve powers of more than one of the operators $s_1, s_2, \cdots, s_\lambda$ is that all these powers be of different orders, and that the larger of two generators involved be raised to a higher power of p than the smaller, and this power have a larger order than the power of the smaller. In particular, no two of these constituents are powers of independent generators whose orders have a ratio which is less than p^2.

As each of the possible sets of I-conjugate operators of G is completely determined by the I-reduced operator which appears in the set, it results that the determination of the number of different sets of I-conjugate operators is equivalent to the determination of the possible number of different I-reduced operators. It should be noted that the number of I-reduced operators depends only upon the orders and the number of the different orders of the independent generators of G. That is, if G has more than one independent generator of the same order, the number of I-reduced operators of G is the same as that of the group having only one of these generators and only one generator whose order is equal to the order of every other independent generator of G.

To determine the number of operators of G which are I-conjugate with a given I-reduced operator T of G, it is convenient to call t-generators all the independent generators of G whose orders are equal to the orders of those independent generators whose powers appear in this I-reduced operator. The remaining independent generators of G will be called s-generators. Let $p^{\beta_1}, p^{\beta_2}, \cdots, p^{\beta_\theta}$ be the indices, in descending order of magnitude, of the various powers of t-generators which appear as constituents of T, and construct a subgroup of G whose independent generators are powers of s-generators which are determined as follows:

All the s-generators of G whose orders exceed the order of the largest

t-generator are raised to powers such that the common order of these powers is equal to that of the p^{β_1} power of this t-generator, and all the s-generators whose orders are smaller than the smallest t-generator are raised to the p^{β_θ} power. Each of the other s-generators of G is raised to the highest power whose index does not exceed the index of the power to which the next larger t-generator is raised to obtain a constituent of T and whose order is not less than the order of the power of the next lower t-generator which appears in T. The powers of the s-generators thus determined constitute a set of independent generators of the subgroup in question.

To obtain all the operators of the set of I-conjugate operators of T, we multiply all the operators of the subgroup noted in the preceding paragraph by the product of the operators of highest orders in the θ subgroups which are separately generated by the p^{β_1}, p^{β_2}, $\cdots$, p^{β_θ} powers respectively of the t-generators of the same order contained in G. As the invariants of each of these θ subgroups are equal to each other, these powers for any particular subgroup are evidently I-conjugate, but the powers for one subgroup are not I-conjugate with the powers in question contained in another of these θ subgroups. In particular, the number of operators in each set of I-conjugate operators of G besides identity is divisible by $p-1$, as results also directly from the fact that an automorphism of an abelian group can be obtained by letting each operator correspond to any given power of itself whose index is prime to the order of the group.

If the order of an abelian group is not a power of a prime number, the number of its sets of I-conjugate operators is evidently the product of the numbers of the sets of I-conjugate operators of its Sylow subgroups. In particular, it may be desirable to emphasize the theorem: *The number of sets of I-conjugate operators in any abelian group is equal to the product of the numbers of the I-reduced operators in its Sylow subgroups for a set of independent generators in which the order of each generator is a power of a prime number.* In this theorem, identity is included among the I-reduced operators of a Sylow subgroup.

For the purpose of illustrating the preceding developments, we shall consider the special abelian group of order p^{10} and of type (6, 3, 1). In addition to identity, the number of I-reduced operators involving a single constituent is 10, the number of those involving two constituents is 11, and the number involving three constituents is 2. Hence this group involves 24 sets of I-conjugate operators including identity. The numbers of I-conjugate operators in these 24 sets are as follows: 1, $p-1$, p^2-p, p^3-p^2, p^5-p^4, p^7-p^6, $p^{10}-p^9$, p^2-p, p^4-p^3, p^7-p^6, p^3-p^2, $(p^2-p)(p-1)$, $(p^3-p^2)(p-1)$, $(p^3-p^2)(p-1)$, $(p^4-p^3)(p-1)$, $(p^4-p^3)(p-1)$, $(p^5-p^4)(p-1)$, $(p^5-p^4)(p-1)$, $(p^7-p^6)(p-1)$, $(p^7-p^6)(p-1)$,

$(p^5-p^7)(p-1)$, $(p^4-p^3)(p-1)$, $(p^4-p^3)(p-1)^2$, $(p^5-p^4)(p-1)^2$. These numbers illustrate the obvious theorem that a necessary and sufficient condition that an operator of an abelian group of order p^m, $p>2$, having no two invariants which are equal to each other, be either a possible independent generator or a power of such a generator is that the number of its I-conjugates be not divisible by $(p-1)^2$.

This theorem is evidently a special case of the theorem that a necessary and sufficient condition that the I-reduced operator of a set of I-conjugate operators involve powers of α operators of a set of independent generators of G is that the number of operators in this set be divisible by $(p-1)^\alpha$ for a general value of p. It should be noted that the number of I-conjugate sets of operators of a group of order p^m depends on the type of this group, but is independent of the value of the prime number p, and that the theorem stated at the close of the preceding paragraph is not affected by the number of independent generators of the same order when G has a general value. For special given values of p, the theorem stated at the opening of the present paragraph is clearly not always valid.

III. Criteria for I-conjugate operators and for certain I-conjugate subgroups

It was noted in the preceding section that there is one and only one I-reduced operator in every complete set of I-conjugate operators of the abelian group G of order p^m, and that the number of constituents in terms of a fixed set of independent generators of G appearing in such an I-reduced operator can be determined from the number of operators involved in the set of I-conjugates to which this I-reduced operator belongs. A necessary and sufficient condition that two operators of G be I-conjugate is that they be I-conjugate with the same I-reduced operators. We proceed to develop another criterion for determining when two operators are I-conjugate.

The cyclic group generated by an I-reduced operator gives rise to a quotient group which is known to be simply isomorphic with a subgroup of G. The s-generators of G and all the t-generators of G with respect to this I-reduced operator except those whose powers actually appear in it can also be used as independent generators of the quotient group in question. To each of the latter t-generators, except the smallest one, there corresponds a generator of this quotient group whose order exceeds the order of all these t-generators whose order is less than that of the t-generator in question.

The quotient groups which correspond to two I-conjugate cyclic subgroups are evidently of the same type. To prove that, conversely, every two cyclic subgroups which give rise to quotient groups of the same type are I-conjugate, it should be noted that when these cyclic groups are replaced by those generated

by the I-reduced operators in the sets of I-conjugate operators to which their generators belong, their largest constituent groups with respect to the set of independent generators of G in question must be generated by the same power of independent generators of the same order, since, otherwise, in one quotient group the number of independent generators, beginning with the largest, whose orders coincide with those of G would differ from the number of the corresponding independent generators of the other quotient group.

If the generators of the cyclic groups in question involve powers of more than one t-generator of G, the second t-generator involved must again be the same for both of these cyclic groups, since the independent generators of the quotient group which corresponds to the first t-generator are of a larger order than the second t-generator, as was noted above. Moreover, the same power of this second t-generator must appear in a generator of each of the two cyclic subgroups in question. As this process may be continued until all the t-generators whose powers appear in the constituents of the cyclic subgroups under consideration have been exhausted, there results the following:

THEOREM. *A necessary and sufficient condition that two operators of any abelian group be I-conjugate is that the cyclic groups generated by these operators give rise to quotient groups which are of the same type.*

It results directly from the preceding theorem that the number of different sets of I-conjugate operators can be determined by counting the number of different types of quotient groups to which cyclic subgroups of G give rise. For instance, the cyclic subgroups of the abelian group of order p^4 and of type (3, 1) clearly give rise to quotient groups of the following types, and of no other types: (3, 1), (3), (2, 1), (2), (1, 1), (1). Hence this group has exactly six sets of I-conjugate operators, including identity. The number of operators in these sets is 1, $p^2 - p$, $p - 1$, $p(p - 1)^2$, $p^2 - p$, $p^4 - p^3$, respectively. All of these operators are either possible independent generators or powers of such generators except those of the fourth set.

As a first step in a proof of the theorem that if two subgroups of the same type give rise to cyclic quotient groups they must be I-conjugate, it will be convenient to consider a necessary and sufficient condition that a subgroup H of G give rise to a cyclic quotient group. If G/H is cyclic, and if as many as possible of the operators of a set of independent generators of G are selected from the operators of H, the remaining operators of this set can be chosen as follows:

As one of the operators any operator s_1 of lowest order contained in a co-set corresponding to any operator of highest order in G/H may be selected. A necessary and sufficient condition that s_1 be the only operator of the set of independent generators in question which does not appear in H is that one of the operators of smallest order in every co-set corresponding to an operator

of G/H be a power of s_1. When this condition is not satisfied, find one of the largest operators of G/H such that a power of s_1 is not one of the smallest operators in the corresponding co-set. Let s_2 be any one of the smallest operators in such a co-set. It is evident that s_2 may then be chosen as a second operator of the set of independent generators in question.

If the powers of s_2 are not operators of lowest order in the co-sets with respect to H in which they appear, we select an operator s_3 of lowest order in the co-set which corresponds to the largest operator of G/H to which such an operator corresponds but is not an operator of lowest order in the co-set. This process is continued until an operator s_α is found such that its powers are operators of lowest order in all the co-sets with respect to H in which they appear. The operators $s_1, s_2, \cdots, s_\alpha$ are then the operators of the set of independent generators in question which do not appear in H. It may be noted that the ratio of the order of any one of these operators and the order of the one which follows it in this sub-set cannot be less than p^2.

For the independent generators of H which are not also independent generators of G we may choose the operators

$$s_1^{p^{\rho_1}} s_2, s_2^{p^{\rho_2}} s_3, \cdots, s_{\alpha-1}^{p^{\rho_{\alpha-1}}} s_\alpha, s_\alpha^{p^{\rho_\alpha}},$$

where $s_x^{p^{\rho_x}}$ $(x = 1, 2, \cdots, \alpha - 1)$ is the inverse of the lowest power of s_x which appears in a co-set with respect to H in which it is not an operator of lowest order, and $s_\alpha^{p^{\rho_\alpha}}$ is the lowest power of s_α which is found in H. The order of $s_x^{p^{\rho_x}}$ must exceed the order of s_{x+1}, since the latter is an operator of lowest order in the co-set in which the former appears and is an independent generator which cannot be replaced by an operator of H. Hence it follows that *a necessary and sufficient condition that a subgroup G give rise to a cyclic quotient group is that the α independent generators $(s_1, s_2, \cdots, s_\alpha)$ of G which cannot be selected from H can be so chosen that*

$$s_\alpha^{p^{\rho_\alpha}} \quad \textit{and} \quad s_x^{p^{\rho_x}} s_{x+1} \qquad (x = 1, 2, \cdots, \alpha - 1)$$

are independent generators of H, where $s_x^{p^{\rho_x}}$ is of a larger order than s_{x+1} and $\rho_x > 0$. It should be noted that each of these independent generators of H is of a larger order than any of the independent generators of G whose powers appear in the succeeding independent generators of H.

By means of the theorem of the preceding paragraph, it is easy to find a necessary and sufficient condition that two subgroups H_1, H_2 of G which give rise to cyclic quotient groups be I-conjugate. It is evident that a necessary condition is that H_1 and H_2 be of the same type. To prove that this is also a sufficient condition, when it is assumed that both of the quotient groups G/H_1 and G/H_2 are cyclic, let $s_1, s_2, \cdots, s_\rho$ and $t_1, t_2, \cdots, t_\rho$ be two sets of independent generators of G which have been so chosen that as many as possible

of these operators are selected from those of H_1 and H_2 respectively. It results that the first operators of each of these sets, arranged in the descending order of magnitude, whose orders exceed the order of the corresponding reduced independent generator of H_1 and H_2, respectively, arranged similarly, must have the same order. The largest independent generator of H_1 which is not also an independent generator of G has the same order as the largest independent generator of H_2 which is not also an independent generator of G, since the order of this independent generator must exceed the order of all the other independent generators of G which are not also independent generators of H_1 or H_2. Hence it results that these independent generators of H_1 and H_2 may be regarded as products of powers of independent generators of G which are of the same order and independent generators of next to the highest order which are found in G but not in the H's.

Just as the operators of H_1 and H_2 which correspond to the largest independent generator of G which is not also an independent generator of H_1 or H_2 can be chosen from the operators of G as $s_1^{p_1} s_2$ was chosen, so the operators which correspond to the next to the largest independent generator of G which is not also an independent generator of H_1 or H_2 can be chosen in the same way as $s_2^{p_2} s_3$ was chosen, whenever not all the independent generators of G save one can be selected from the operators of H_1. Since these arguments apply to these successive independent generators, it results that *every two subgroups of G which are both of the same type and give rise to cyclic quotient groups are I-conjugate.*

If H_1 and H_2 are two subgroups of the same type which give rise to two quotient groups of the same type, it does not necessarily follow that H_1 and H_2 are I-conjugate, as may be seen by considering the group G of order p^9 and of type (5, 3, 1). If s_1, s_2, s_3 represent the three generators of G of orders p^5, p^3, and p respectively, and if $s_1^p s_2$, s_2^p and s_1^p, $s_2^p s_3$ are the independent generators of H_1 and H_2 respectively, it results that the two quotient groups G/H_1 and G/H_2 are of type (2, 1), and the two groups H_1, H_2 are of type (4, 2). The latter groups cannot be I-conjugate, since the operators of the highest order in the latter are powers of operators of highest order in G, but this is not the case as regards the operators of highest order of the former subgroup.

It may be noted that the sum of the number of independent generators of a subgroup of G plus the number of the independent generators of the quotient group corresponding to this subgroup is equal to the number of independent generators of G whose common order is p increased by a number which may vary from the number of independent generators of G whose orders exceed p to twice this number, but can have no other value. Both of the limiting values can evidently be actually attained, and the fact that this sum can have no other values results from the theorem that a quotient group of an abelian

group is always simply isomorphic with a subgroup of this group, and that the independent generators of a subgroup which gives rise to a cyclic quotient group can be selected as noted above.

It was noted above that when two subgroups of the same type give rise to cyclic quotient groups they must be I-conjugate, and when two cyclic subgroups give rise to quotient groups of the same type they are also I-conjugate. The other extreme cases are when two subgroups of the same type give rise to quotient groups of type $(1, 1, 1, \cdots)$ and when two subgroups of type $(1, 1, 1, \cdots)$ give rise to quotient groups of the same type. In each of these two cases the two subgroups in question are again I-conjugate. In the special case when a subgroup gives rise to a quotient group of type $(1, 1, 1, \cdots)$ which involves as many invariants as G itself, the subgroup is characteristic, being the ϕ-subgroup of G. In this special case, the subgroup is completely determined by the type of the quotient group to which it gives rise.

Every subgroup of G which gives rise to a quotient group of type $(1, 1, 1, \cdots)$ must include the ϕ-subgroup of G. If the ϕ-quotient group is of order p^α and a subgroup H gives rise to a quotient group of order p^β and of type $(1, 1, 1, \cdots)$, it results that exactly $\alpha - \beta$ of the independent generators of G are found in H, while the pth power of each of the other independent generators of G is found in this subgroup. From this it results directly that if two subgroups of the same type give rise to quotient groups of type $(1, 1, 1, \cdots)$ these subgroups must be I-conjugate. The number of the characteristic subgroups which give rise to quotient groups of type $(1, 1, 1, \cdots)$ is evidently equal to the number of the different orders of reduced independent generators of G. As every operator of order p found in G is a power of a possible independent generator of G, it results that when two subgroups of type $(1, 1, 1, \cdots)$ give rise to quotient groups of the same type their independent generators can be so chosen that they are powers of independent generators of G which are of the same orders. Hence these subgroups are I-conjugate.

UNIVERSITY OF ILLINOIS,
URBANA, ILL.

GENERALIZED LIMITS IN GENERAL ANALYSIS, FIRST PAPER*

BY

CHARLES N. MOORE

The analogies that exist between infinite series and infinite integrals are well known and have frequently served to indicate the extension of a theorem or a method from one of these domains of investigation to the other. According to a principle of generalization that has been formulated by E. H. Moore, the presence of such analogies implies the existence of a general theory which includes the central features of both the special theories.† It is the purpose of the present paper to develop the fundamental principles of that section of this general theory which contains as particular instances the theories of Cesàro and Hölder summability of divergent series and divergent integrals. Furthermore, the usefulness of the theory will be illustrated by proving a general theorem in it which includes as special cases the Knopp-Schnee-Ford theorem‡ with regard to the equivalence of the Cesàro and Hölder means for summing divergent series, an analogous theorem due to Landau § concerning divergent integrals, and a further new theorem with regard to the equivalence of certain generalized derivatives.

The general theorem just mentioned can be extended to the case of multiple limits so as to include other new theorems, analogous to those referred to above, with regard to multiple series, multiple integrals, and partial derivatives. This extension, however, involves formulas that are considerably more complicated than in the case of simple limits. I shall therefore reserve it for a second paper, as I wish to avoid algebraic complexity in this first presentation of the general theory.

Following the terminology introduced by E. H. Moore,‖ we indicate the basis of our general theory as follows:

$$(\mathfrak{A};\ \mathfrak{P};\ \mathfrak{S};\ \mathfrak{G}^{\text{on } \mathfrak{S} \text{ to } \mathfrak{A}};\ \mathfrak{H}^{\text{on } \mathfrak{S} \text{ to } \mathfrak{A}};\ \mathfrak{F}^{\text{on } \mathfrak{S} \text{ to } \mathfrak{A}};\ \phi_0;\ J^{\text{on } \mathfrak{W} \text{ to } \mathfrak{C};\ \text{on } \mathfrak{P} \text{ to } \mathfrak{F}})$$

* Presented to the Society, December 28, 1918.

† Cf. E. H. Moore, *Introduction to a Form of General Analysis*, The New Haven Mathematical Colloquium, Yale University Press, 1910, p. 1.

‡ That the existence of the Hölder limit implies the existence and equality of the Cesàro limit of the same order was first proved by Knopp; cf. his Inauguraldissertation, *Grenzwerte von Reihen bei der Annäherung an die Konvergenzgrenze*, Berlin, 1907. The converse theorem was established independently by Schnee and W. B. Ford; cf. Mathematische Annalen, vol. 67 (1909), pp. 110–125, and American Journal of Mathematics, vol. 32 (1910), pp. 315–326.

§ See Leipziger Berichte, vol. 65 (1913), pp. 131–138.

‖ Cf. his two papers: *On the foundations of the theory of linear integral equations*, Bulletin

where $\mathfrak{A} \equiv [a]$ denotes the class of all real numbers a, $\mathfrak{P} \equiv [p]$ denotes a class of elements p, and $\mathfrak{S} \equiv [\sigma]$ denotes a class of sets σ of elements p of the range $\mathfrak{P}$; $\mathfrak{G} \equiv [\gamma]$, $\mathfrak{H} \equiv [\eta]$, and $\mathfrak{F} \equiv [\phi]$ are three classes of functions γ, η, and ϕ respectively on $\mathfrak{S}$ to $\mathfrak{A}$ (we shall restrict ourselves throughout to the consideration of single-valued functions); ϕ_0 is a special function ϕ of the class $\mathfrak{F}$; and J is a function on $\mathfrak{G}$ to $\mathfrak{H}$ and on $\mathfrak{H}$ to $\mathfrak{F}$, that is a functional transformation turning a function of the class $\mathfrak{G}$ into a function of the class $\mathfrak{H}$ or a function of the class $\mathfrak{H}$ into a function of the class $\mathfrak{F}$, denoted by $J\gamma$ or $J\eta$.

In order to make clear the relationship of our general theorem to the two special theorems referred to above we will indicate here what the general basis reduces to in the particular instances III and IV:

$$\mathfrak{P}^{\mathrm{III}} \equiv [\text{all } n = 1, 2, 3, \cdots]; \qquad \mathfrak{S} \equiv [\sigma_n \equiv (1, 2, \cdots, n) \mid n];$$

$$\mathfrak{G} \equiv \mathfrak{H} \equiv \mathfrak{F} \equiv [\text{all } \gamma, \eta, \phi^{\text{on } \mathfrak{S} \text{ to } \mathfrak{A}}]; \qquad \phi_0(\sigma_n) = n \ (n);$$

$$(J\gamma)(\sigma_n) = \gamma(\sigma_1) + \gamma(\sigma_2) + \cdots + \gamma(\sigma_n) \ (n);$$

$$(J\eta)(\sigma_n) = \eta(\sigma_1) + \eta(\sigma_2) + \cdots + \eta(\sigma_n) \ (n);$$

$$\mathfrak{P}^{\mathrm{IV}} \equiv [\text{all } a > 0]; \qquad \mathfrak{S} \equiv [\sigma \equiv (\text{all } x \text{ such that } 0 < x \leqq a) \ (a > 0)];$$

$$\mathfrak{G} \equiv [\text{all functions that are finite and integrable (Lebesgue) on every finite interval } (0 < x \leqq a)];$$

$$\mathfrak{H} \equiv \left[\text{all } \eta = \left(\int_{\ }^{x} \gamma \mid x > 0\right)\right]; \qquad \mathfrak{F} \equiv \left[\text{all } \phi = \left(\int_0^x \eta \mid x > 0\right)\right];$$

$$\phi_0(\sigma_a) = a \ \ (a); \quad (J\gamma)(\sigma_a) = \int_0^a \gamma(\gamma a); \quad (J\eta)(\sigma_a) = \int_0^x \eta(\eta a).$$

We next proceed to make certain postulates with regard to the nature of the elements in our basis, readily seen to be verified in the specific instances indicated. Thus we require the class $\mathfrak{S}$ to have the following properties:

(U) Either corresponding to every σ' there exists a least common superclass of classes $\sigma < \sigma'$, or there exists a σ_0 such that for every $\sigma' > \sigma_0$ there exists a least common superclass of classes $\sigma < \sigma'$. In both cases the least common superclass is itself a σ.

(A) Corresponding to every σ' there exists a least common subclass of classes $\sigma > \sigma'$, this least common subclass being itself a σ.

In the typical instances in view in the formation of this general theory, of the two alternatives in (U) one holds and the other does not hold; however, it is not assumed that this disjunction between the two alternatives shall be

of the American Mathematical Society, vol. 18 (1912), pp. 334–362; *On the fundamental functional operation of a general theory of linear integral equations*, Proceedings of the Fifth International Congress of Mathematicians, Cambridge, 1913, pp. 230–255.

presupposed. In order to avoid notational reference to these alternatives, it is convenient to introduce a property $-$ of sets σ of $\mathfrak{S}$; if the first alternative holds, every σ of $\mathfrak{S}$ has the property $-$, in notation $\bar{\sigma}$; if the first alternative does not hold, the sets $\bar{\sigma}$ are the sets $\sigma > \sigma_0$; further for brevity we introduce a property $\cdot$ (the negation of $-$); thus every σ is a $\bar{\sigma}$ or a $\dot{\sigma}$.

We now define

$$\sigma'_{-1} \equiv [\text{the least common superclass of classes } \sigma < \sigma'] \quad (\bar{\sigma'});$$

$$\sigma'_1 \equiv [\text{the least common subclass of classes } \sigma > \sigma'];$$

$$\sigma'_{n+1} \equiv [\text{the least common subclass of classes } \sigma > \sigma'_n] \ (\sigma \mid n = 1, 2, \cdots).$$

We then postulate

(R) Corresponding to every σ', σ'_1 is a $\bar{\sigma}$ and there exists $(\sigma'_1)_{-1} = \sigma'$.

We next define the notation $\sigma' < \sigma''$, $\sigma'' > \sigma'$, to mean that σ'' contains all the elements of σ' and at least one element not found in σ'. We are then ready to formulate three limit definitions which are based on the fundamental definition of limit in General Analysis given by E. H. Moore.*

For a given θ on $\mathfrak{S}$ to $\mathfrak{A}$, a given a, and a given σ' such that there exists $\sigma < \sigma'$, we shall write

$$\lim_{\sigma \mid \sigma < \sigma'} \theta(\sigma) = a$$

in the case that, corresponding to an arbitrary positive number e, there exists a $\sigma_e < \sigma'$ such that for every σ having the property $\sigma_e \leqq \sigma < \sigma'$, $|\theta(\sigma) - a| < e$.

For any function θ on $\mathfrak{S}$ to $\mathfrak{A}$ we shall say that $\theta(\sigma)$ approaches a limit as to σ if corresponding to every positive e we can find a σ_e such that for every $\sigma > \sigma_e$ we have $|\theta(\sigma) - a| < e$.

For any function θ on $\mathfrak{S}$ to $\mathfrak{A}$ we shall mean by the notation

$$\lim_{\sigma} \theta(\sigma) = \infty$$

that for every positive e there exists a σ_e such that for every $\sigma > \sigma_e$, $\theta(\sigma) > e$.

We define the notation

$$(D\theta)(\sigma) = \alpha(\sigma)$$

with regard to every θ, α on $\mathfrak{S}$ to $\mathfrak{A}$, to mean that

$$(1) \qquad \lim_{\sigma \mid \sigma < \sigma'} \frac{\theta(\sigma') - \theta(\sigma)}{\phi_0(\sigma') - \phi_0(\sigma)} = \alpha(\sigma') \quad (\bar{\sigma'}), \qquad \theta(\sigma) = \alpha(\sigma) \quad (\dot{\sigma}).$$

* Cf. Proceedings of the National Academy of Sciences, vol. 1 (1915), pp. 628–632.

We require the class $\mathfrak{G}$ to have the linear property

(L) as defined by E. H. Moore,*

and the property (P) defined by

(P) The product $\gamma_1(\sigma)\cdot\gamma_2(\sigma)$ is a function of the class $\mathfrak{G}$.

It will then follow that the product $\gamma_1(\sigma)\cdot\gamma_2(\sigma)\cdot\cdots\cdot\gamma_n(\sigma)$ is a function of the class $\mathfrak{G}$.

We require the class $\mathfrak{H}$ to have the properties (L) and (P) and the further property of being a subclass of the class $\mathfrak{G}$, which property we designate as S_G. We further postulate for the class $\mathfrak{H}$ the property (B) defined by

(B) If $\lim_{\sigma} \eta(\sigma)$ exists and is equal to a, then $|\eta(\sigma)| < a_1 \quad (\sigma)$.

We require the class $\mathfrak{F}$ to have the properties (L) and (P) and the further property of being a subclass of the class $\mathfrak{H}$, which property we designate as S_H. Hence J is also on $\mathfrak{F}$ to $\mathfrak{F}$. We also postulate for the class $\mathfrak{F}$ the property (Δ) defined by

(Δ) There exists $D\phi \equiv [(D\phi)(\sigma)|\sigma]$, a function of the class $\mathfrak{H}$.

We now define

$$(2)\quad \bar{\phi}(\sigma) \equiv \phi(\sigma_{-1}) \quad (\bar{\sigma}), \qquad \bar{\phi}(\sigma) \equiv 0 \quad (\dot{\sigma}); \qquad \dot{\phi}(\sigma) \equiv \phi(\sigma_1) \quad (\sigma);$$

and with regard to the functions $\bar{\phi}$ and $\dot{\phi}$ we postulate

(F) All functions $\bar{\phi}$ and $\dot{\phi}$ are of the class $\mathfrak{F}$.

We further postulate with regard to the class $\mathfrak{F}$

(C) For every ϕ and every $\bar{\sigma}'$ there exists $\lim_{\sigma'\sigma<\sigma'} \phi(\sigma) = \bar{\phi}(\bar{\sigma}')$.

For the operation J we postulate the following properties:

(M_1) If $a_1 < \gamma_1 < a_2$, $0 \leqq \gamma_2$, then

$$a_1(J\gamma_2)(\sigma) \leqq (J[\gamma_1\gamma_2])(\sigma) \leqq a_2(J\gamma_2)(\sigma) \quad (\sigma),$$

(M_2) If for every $\sigma > \sigma'$, $a_1 < \gamma_1 < a_2$, $0 \leqq \gamma_2$, then for $\sigma'' > \sigma'$

$$a_1[(J\gamma_2)(\sigma'') - (J\gamma_2)(\sigma')] \leqq (J[\gamma_1\gamma_2])(\sigma'') - (J[\gamma_1\gamma_2])(\sigma') \leqq a_2[(J\gamma_2)(\sigma'') - (J\gamma_2)(\sigma')],$$

(I_D) For every η and every σ there exists $(D(J\eta))(\sigma) = \eta(\sigma)$,

(I_J) For every ϕ and every σ there exists $(J(D\phi))(\sigma) = \phi(\sigma)$.

We next introduce for the sake of brevity the following notations:

$$(3)\quad \phi_{0n}(\sigma) = \phi_0(\sigma)\cdot\phi_0(\sigma_1)\cdot\phi_0(\sigma_2)\cdot\cdots\cdot\phi_0(\sigma_{n-1}) \quad (n > 1),$$

$$\phi_{01}(\sigma) = \phi_0(\sigma).$$

* Loc. cit.

We then postulate with regard to ϕ_0

(I) ϕ_0 is a positive increasing function of σ,

(II) $\dfrac{1}{\phi_{0n}(\sigma)}\,(J^n\eta)(\sigma)$ as function of σ is of the class $\mathfrak{F}\quad(n)$,

the symbol $J^n\eta$ indicating that the operation J has been repeated n times,

(III) $\phi_0(\sigma_n)$ as function of σ is of the class $\mathfrak{F}\quad(n)$,

(IV) $\lim_{\sigma}\phi_{0n}(\sigma)=\infty\quad(n)$,

(V) $[\phi_0(\sigma_1)-\phi_0(\sigma)]$ is constant for all σ,

(VI) $\phi_{0n}(\dot\sigma)=n\phi_{0,\,n-1}(\dot\sigma)\quad(n>1),\qquad \phi_{01}(\dot\sigma)=\phi_0(\dot\sigma)=1$.

We have then as the foundation of our theory:

$$\sum\equiv(\mathfrak{A};\ \mathfrak{P};\ \mathfrak{Z}^{UAR};\ \mathfrak{G}^{\text{on }\mathfrak{Z}\text{ to }\mathfrak{A}\cdot LP};\ \mathfrak{H}^{\text{on }\mathfrak{Z}\text{ to }\mathfrak{A}\cdot LPS_GB};$$
$$\mathfrak{F}^{\text{on }\mathfrak{Z}\text{ to }\mathfrak{A}\cdot LPS_HC\mathfrak{A}};\ \phi_0^{\mathfrak{A}\cdot \text{I II III IV V VI}};\ \dot\phi^{\dot{\mathfrak{Z}}};\ \dot{\dot\phi}^{\dot{\mathfrak{Z}}};$$
$$J^{\text{on }\mathfrak{G}\text{ to }\mathfrak{H}\cdot\text{on }\mathfrak{H}\text{ to }\mathfrak{F}\cdot M_1M_2I_DI_J}).$$

We will now prove that the operation J, when applied to the class $\mathfrak{H}$, has the linear property (L). Let us set

$$(J\eta_n)(\sigma)=\phi_n(\sigma)\qquad(n=1,2,\cdots,i).$$

It follows from the definition of D and I_D that

$$(D(a_1\phi_1+a_2\phi_2+\cdots+a_i\phi_i))(\sigma)$$
$$=a_1(D\phi_1)(\sigma)+a_2(D\phi_2)(\sigma)+\cdots+a_i(D\phi_i)(\sigma)$$
$$=a_1\eta_1+a_2\eta_2+\cdots+a_i\eta_i.$$

Applying J to both sides and making use of I_J, we have

$$(J(a_1\eta_1+a_2\eta_2+\cdots+a_i\eta_i)(\sigma)$$
$$=a_1(J\eta_1)(\sigma)+a_2(J\eta_2)(\sigma)+\cdots+a_i(J\eta_i)(\sigma).$$

We will now prove two properties of the operations D and J as applied to the class $\mathfrak{F}$. We have for $\sigma\neq\sigma'$

$$\frac{\phi_1(\sigma')\phi_2(\sigma')-\phi_1(\sigma)\phi_2(\sigma)}{\phi_0(\sigma')-\phi_0(\sigma)}$$
$$=\phi_2(\sigma')\frac{\phi_1(\sigma')-\phi_1(\sigma)}{\phi_0(\sigma')-\phi_0(\sigma)}+\phi_1(\sigma)\frac{\phi_2(\sigma')-\phi_2(\sigma)}{\phi_0(\sigma')-\phi_0(\sigma)}.$$

By virtue of properties (C) and (Δ) of class $\mathfrak{F}$ the right side of the above

equation approaches a limit as to $\sigma | \sigma < \sigma'$, for every $\bar{\sigma}'$. Hence so also does the left side, and we obtain the formula

$$(4) \qquad (D[\phi_1 \phi_2])(\sigma) = \phi_2(\sigma)(D\phi_1)(\sigma) + \bar{\phi}_1(\sigma)(D\phi_2)(\sigma)$$

for $\sigma = \bar{\sigma}$. In view of (1) and (2) equation (4) obviously holds for $\sigma = \dot{\sigma}$.

By virtue of postulates (Δ) and (F) and the properties (P) and (L) of the class $\mathfrak{H}$ we may apply the operation J to equation (4). On doing so we obtain, in view of I_J and the linear property of J established above,

$$(5) \qquad (J[\phi_2(D\phi_1)])(\sigma) = \phi_1(\sigma)\phi_2(\sigma) - (J[\bar{\phi}_1(D\phi_2)])(\sigma).$$

Equation (4) includes as special cases the formulas for differentiation of a product and forming the first difference of a product. Equation (5) includes as special cases the formulas for integration by parts and partial summation.

We shall next prove two further properties of the special function ϕ_0. The first of these properties is the following:

$$(\text{VII}) \qquad (D\phi_{0n})(\sigma) = n\phi_{0,\,n-1}(\sigma) \quad (\sigma, n > 1), \qquad (D\phi_{01})(\sigma) = 1 \quad (\sigma).$$

We have, in view of the definition of D, (V), (R), (C), and (2),

$$(6) \qquad \begin{aligned} (D\phi_{02})(\sigma') &= \lim_{\sigma | \sigma < \sigma'} \left[\phi_0(\sigma') \frac{\phi_0(\sigma_1') - \phi_0(\sigma_1)}{\phi_0(\sigma') - \phi_0(\sigma)} + \phi_0(\sigma_1) \frac{\phi_0(\sigma') - \phi_0(\sigma)}{\phi_0(\sigma') - \phi_0(\sigma)} \right] \\ &= \lim_{\sigma | \sigma < \sigma'} [\phi_0(\sigma') + \phi_0(\sigma_1)] = 2\phi_0(\sigma') \quad (\bar{\sigma}'). \end{aligned}$$

We now assume

$$(D\phi_{0i})(\sigma) = i\phi_{0,\,i-1}(\sigma) \qquad (i \geqq 2, \bar{\sigma}).$$

Then, making use of (4), we have

$$(7) \qquad \begin{aligned} (D\phi_{0,\,i+1})(\sigma) &= (D[\phi_0 \dot{\phi}_{0i}])(\sigma) \\ &= \phi_0(\sigma)(D\phi_{0i})(\sigma_1) + \phi_{0i}(\sigma)(D\phi_0)(\sigma) \\ &= \phi_0(\sigma)[i\phi_{0,\,i-1}(\sigma_1)] + \phi_{0i}(\sigma) = (i+1)\phi_{0i}(\sigma) \quad (\bar{\sigma}). \end{aligned}$$

Hence, if (VII) holds for $\bar{\sigma}$, $n = i$ $(i \geqq 2)$, it will hold for $\bar{\sigma}$, $n = i + 1$. Combining this fact with equation (6), we infer that (VII) holds for $\bar{\sigma}$, $n \geqq 2$; for $\dot{\sigma}$, $n = 1$ it is an obvious consequence of the definition of D. For $\dot{\sigma}$, n it follows at once from the definition of D and (VI).

We now introduce the following notation:

$$(8) \qquad \psi_n(\sigma) = \frac{1}{\phi_{0,\,n-1}(\sigma)} \qquad (n > 1).$$

The second property of ϕ_0 that we wish to prove may then be stated as follows:

$$(\text{VIII}) \qquad (D\psi_n)(\sigma) = -\frac{n-1}{\phi_{0n}(\sigma_{-1})} \qquad (\bar{\sigma}, n > 1).$$

We have from the definition of D, (R), (C), and (2)

$$(9)\qquad \begin{aligned}(D\psi_2)(\sigma') &= \left(D\frac{1}{\phi_{01}}\right)(\sigma') = \lim_{\sigma|\sigma<\sigma'} \frac{\dfrac{1}{\phi_0(\sigma')} - \dfrac{1}{\phi_0(\sigma)}}{\phi_0(\sigma') - \phi_0(\sigma)} \\ &= \lim_{\sigma|\sigma<\sigma'} \frac{-1}{\phi_0(\sigma')\phi_0(\sigma)} = \frac{-1}{\phi_0(\sigma'_{-1})\phi_0(\sigma')} = -\frac{1}{\phi_{02}(\sigma'_{-1})} \quad (\bar{\sigma}).\end{aligned}$$

Then, assuming (VIII) for $n = i$, we have from (4) and (9)

$$(10)\qquad \begin{aligned}(D\psi_{i+1})(\sigma) &= \left(D\left[\frac{1}{\phi_0}\psi_i\right]\right)(\sigma_1) \\ &= -\frac{1}{\phi_{02}(\sigma_{-1})}\cdot\frac{1}{\phi_{0,\,i-1}(\sigma_1)} + \frac{1}{\phi_0(\sigma_{-1})}\cdot\frac{-(i-1)}{\phi_{0i}(\sigma)} \\ &= -\frac{i}{\phi_{0,\,i+1}(\sigma_{-1})} \quad (\bar{\sigma}).\end{aligned}$$

From (9) and (10) the proof by induction of formula (VIII) may readily be completed.

We are now ready to define the two generalized limits with which we shall be concerned. Given any function $\eta(\sigma)$, we set

$$(11)\qquad (C_n\eta)(\sigma) \equiv [n!/\phi_{0n}(\sigma)](J^n\eta)(\sigma) \quad (n),$$

$$(12)\qquad (M\eta)(\sigma) \equiv [1/\phi_0(\sigma)](J\eta)(\sigma),$$

$$(13)\qquad (H_n\eta)(\sigma) \equiv (M^n\eta)(\sigma) \quad (n),$$

where $\phi_{0n}(\sigma)$ is defined as in equation (3) and C and H are used, as is customary, to connote Cesàro and Hölder. If for a fixed n $\lim_\sigma (C_n\eta)(\sigma)$ exists, we define this limit as the generalized limit of type (C_n) for $\eta(\sigma)$. If $\lim_\sigma (H_n\eta)(\sigma)$ exists, we define this limit as the generalized limit of type (Hn) for $\eta(\sigma)$.

We shall prove the equivalence of these two generalized limits. We begin by proving some lemmas.

LEMMA 1. *If we represent by E the identical functional operation $E\theta = \theta(\theta)$, we have the identity*

$$(14)\qquad \left(\left(\frac{n-1}{n}M + \frac{1}{n}E\right)(C_n\eta)\right)(\sigma) = (M(C_{n-1}\eta))(\sigma) \quad (n),$$

where for the sake of uniformity we have set $(C_0\eta)(\sigma) = \eta(\sigma)$.

We have from the definition of $(C_n\eta)(\sigma)$

$$(15)\qquad \begin{aligned}\phi_0(\sigma)\cdot(C_n\eta)(\sigma) = {} & \frac{n!}{\phi_0(\sigma_1)\phi_0(\sigma_2)\cdots\phi_0(\sigma_{n-1})} \\ & \times\left(J\left[\frac{1}{(n-1)!}\phi_{0,\,n-1}(C_{n-1}\eta)\right]\right)(\sigma) \quad (n>1).\end{aligned}$$

Applying the operation D to this equation, and making use of (4), (VIII), the property (I_D) of J, and the property (R) of $\mathfrak{Z}$, we get

$$(D[\phi_0(C_n\eta)])(\sigma) = -\frac{(n-1)(n!)}{\phi_0(\sigma)\phi_0(\sigma_1)\cdots\phi_0(\sigma_{n-1})} \times\left(J\left[\frac{1}{(n-1)!}\phi_{0,\,n-1}(C_{n-1}\eta)\right]\right)(\sigma) + n(C_{n-1}\eta)(\sigma).$$

Applying the operation J to this equation, and making use of property (I_J) and equation (15), we obtain

$$\phi_0(\sigma)(C_n\eta)(\sigma) = -(n-1)(J(C_n\eta))(\sigma) + n(J(C_{n-1}\eta))(\sigma).$$

Transposing the first term on the right-hand side and dividing through by $n\phi_0(\sigma)$, we have finally the identity (14) for $n>1$. For $n=1$ it is an obvious consequence of (11).

Before stating the next lemma we need to introduce the following notation:

$$\begin{aligned}(16)\qquad &\phi_n(\sigma) = \phi_0(\sigma)\phi_0(\sigma_1)\cdots\phi_0(\sigma_{n-2})\phi(\sigma) \quad (n>2),\\ &\phi_2(\sigma) = \phi_0(\sigma)\phi(\sigma), \qquad \phi_1(\sigma) = \phi(\sigma).\end{aligned}$$

LEMMA 2. *If* $\lim_\sigma \phi(\sigma)$ *exists and is equal to* a *and* $|\phi(\sigma)| < a_1$ *for every* σ, *then* $\lim_\sigma [\phi_{0n}(\sigma)]^{-1}(J\phi_n)(\sigma)$ *will exist and be equal to* a/n *and we shall have*

$$|[\phi_{0n}(\sigma)]^{-1}(J\phi_n)(\sigma)| < \frac{a_1}{n} \qquad (\sigma).$$

Given a positive e, we choose σ'_e so that $a - e/4 < \phi(\sigma) < a + e/4$ for $\sigma > \sigma'_e$. We have

$$\begin{aligned}(17)\qquad [\phi_{0n}(\sigma)]^{-1}(J\phi_n)(\sigma) &= [\phi_{0n}(\sigma)]^{-1}(J\phi_n)(\sigma'_e)\\ &\quad + [\phi_{0n}(\sigma)]^{-1}[(J\phi_n)(\sigma) - (J\phi_n)(\sigma'_e)].\end{aligned}$$

In view of (16), (3), and (VII) we have the relationship

$$(18)\qquad (J\phi_n)(\sigma) = \left(J\left[\frac{1}{n}(D\phi_{0n})\phi\right]\right)(\sigma).$$

Making use of (18) and postulates M_2 and I_J, we see that the second term on the right-hand side of (17) lies between

$$\frac{1}{n}\left(a-\frac{e}{4}\right)\cdot\left[1-\frac{\phi_{0n}(\sigma'_e)}{\phi_{0n}(\sigma)}\right] \quad\text{and}\quad \frac{1}{n}\left(a+\frac{e}{4}\right)\cdot\left[1-\frac{\phi_{0n}(\sigma'_e)}{\phi_{0n}(\sigma)}\right].$$

We see from (IV) that for a proper choice of $\sigma''_e > \sigma'_e$, each of the above expressions differs from a/n by a quantity less in absolute value than $\frac{1}{2}e$ for all $\sigma > \sigma''_e$.

The first term on the right side of (17) is seen from (18) and (M_1) to be less in absolute value than

$$\frac{a_1}{n}\left|\frac{\phi_{0n}(\sigma'_\epsilon)}{\phi_{0n}(\sigma)}\right| .$$

It follows from (IV) that we can choose $\sigma''_\epsilon > \sigma'_\epsilon$ so as to make this latter expression less in absolute value than $\frac{1}{2}e$ for $\sigma > \sigma''_\epsilon$.

If now we choose for σ_ϵ the greater of σ''_ϵ and σ'''_ϵ, it follows from (17) that for $\sigma > \sigma_\epsilon$, $|[\phi_{0n}(\sigma)]^{-1}(J\phi_n)(\sigma) - a/n| < e$. The first statement in our conclusion is therefore established. We may readily infer that the second statement holds also if we note that in view of (18) and postulates M_1 and I_J, we have

$$-\frac{a_1}{n} < [\phi_{0n}(\sigma)]^{-1}(J\phi_n)(\sigma) < \frac{a_1}{n} \qquad (\sigma).$$

Let us set

$$\phi'(\sigma) = \frac{1}{n}\phi(\sigma) + \frac{n-1}{n}\cdot\frac{1}{\phi_0(\sigma)}(J\phi)(\sigma). \tag{19}$$

We shall then prove

LEMMA 3. *If $\lim_\sigma \phi'(\sigma)$ exists and is equal to a and $|\phi'(\sigma)| < a_1$ for every σ, then $\lim_\sigma \phi(\sigma)$ will exist and be equal to a and we shall have $|\phi(\sigma)| < a_2$ for every σ.*

We define $\phi'_n(\sigma)$ in a manner analogous to that in which $\phi_n(\sigma)$ is defined by (16). Then multiplying (19) by $n\phi_0(\sigma)\phi_0(\sigma_1)\cdots\phi_0(\sigma_{n-2})$ or by $2\phi_0(\sigma)$ according as $n > 2$ or $n = 2$, and making use of (VII) and (3), we have

$$n\phi'_n(\sigma) = \phi_n(\sigma) + [(D\phi_{0,\,n-1})(\sigma_1)]\cdot[(J\phi)(\sigma)] \qquad (n \geqq 2).$$

Applying the operation J to this equation and making use of (5), (I_D), and (I_J), we obtain

$$n(J\phi'_n)(\sigma) = (J\phi_n)(\sigma) + \phi_{0,\,n-1}(\sigma_1)\cdot(J\phi)(\sigma) - (J\phi_n)(\sigma)$$
$$= \phi_{0,\,n-1}(\sigma_1)\cdot(J\phi)(\sigma) \quad (n \geqq 2).$$

Combining the above equation with (19), we have

$$\phi(\sigma) = n\phi'(\sigma) - \frac{n(n-1)}{\phi_0(\sigma)\phi_0(\sigma_1)\cdots\phi_0(\sigma_{n-1})}(J\phi'_n)(\sigma) \qquad (n \geqq 2).$$

Applying Lemma 2 we see that the second term of the right side of this equation approaches $-(n-1)a$ as a limit and remains finite for all σ. Hence our lemma is proved for the case $n \geqq 2$. For $n = 1$ it is an obvious consequence of (19).

Let us set

$$\left(\left(\frac{n-1}{n}M+\frac{1}{n}E\right)\gamma\right)(\sigma)=(S_n\gamma)(\sigma)\qquad(n).$$

Noting that S_n and M are interchangeable operations, we have, from successive applications of (14),

$$(H_1\eta)(\sigma)=(M\eta)(\sigma)=(M(C_0\eta))(\sigma)=(S_1(C_1\eta))(\sigma),$$

$$\begin{aligned}(H_2\eta)(\sigma)&=(M(H_1\eta))(\sigma)=(M(S_1(C_1\eta)))(\sigma)\\&=(S_1(M(C_1\eta)))(\sigma)=(S_1(S_2(C_2\eta)))(\sigma),\end{aligned}$$

. . . . . .

. . . . . .

$$\begin{aligned}(H_n\eta)(\sigma)&=(M(H_{n-1}\eta))(\sigma)\\&=(M(S_1(S_2(S_3\cdots(S_{n-1}(C_{n-1}\eta))\cdots))))(\sigma)\\&=(S_1(S_2(S_3\cdots(S_{n-1}(M(C_{n-1}\eta)))\cdots)))(\sigma)\\&=(S_1(S_2(S_3\cdots(S_{n-1}(S_n(C_n\eta)))\cdots)))(\sigma).\end{aligned}$$

We are now ready to prove the general equivalence theorem:

THEOREM. *If* $\lim_\sigma(C_n\eta)(\sigma)$ *exists and is equal to* a, *then* $\lim_\sigma(H_n\eta)(\sigma)$ *will exist and be equal to* a, *and conversely.*

From the last equation above, property (B), and successive applications of Lemma 2 for the case $n=1$, we infer that the existence of $\lim_\sigma(C_n\eta)(\sigma)=a$ implies the existence of $\lim_\sigma(H_n\eta)(\sigma)=a$, for every n. From the same equation, property (B), and successive applications of Lemma 3 we draw the converse conclusion. Thus our theorem is established.

UNIVERSITY OF CINCINNATI,
CINCINNATI, OHIO.

ANHARMONIC POLYNOMIAL GENERALIZATIONS OF THE NUMBERS OF BERNOULLI AND EULER*

BY

E. T. BELL

We consider twelve infinite systems of polynomials in z which for $z = 1$ degenerate either to the numbers of Bernoulli or Euler, or to others simply dependent upon these. The first part proceeds from the definition of anharmonic polynomials to the specific twelve systems discussed; the second presents an adaptation of the symbolic calculus of Blissard and Lucas in sufficient detail for rapidly developing a simple isomorphism between the algebra of the polynomials and that of the twelve elliptic functions sn, cn, ns, nc, sc, $\cdots$ of Glaisher, and the third contains a short selection from the simpler algebraic and congruential relations between the polynomials. Incidentally there is pointed out in the second part a new interpretation of Kronecker's work on certain symmetric functions and their connections with Bernoulli's numbers. Owing to the length of the paper the development stops short of the quadratic transformation of the polynomials which corresponds to the transformation of the second order in elliptic functions, but the material given is a necessary foundation for all higher transformations. For the same reason only prime moduli are considered in the congruences, although the case in which the modulus is a power of a prime can be treated in essentially the same way, but at greater length. All references are at the end of the paper.

I. Anharmonic polynomials

1. With each of the substitutions σ' of the cross ratio group on z,

$$1 = (z, z), \qquad \alpha' = (z, 1/(1-z)), \qquad \beta' = (z, (z-1)/z),$$
$$\gamma' = (z, 1/z), \quad \delta' = (z, 1-z), \qquad \epsilon' = (z, z/(z-1)),$$

associate a multiplier σ'' as follows,

$$1, \qquad \alpha'' = (z-1)^n, \qquad \beta'' = (-z)^n,$$
$$\gamma'' = z^n, \qquad \delta'' = (-1)^n, \qquad \epsilon'' = (1-z)^n,$$

n being an integer $\geqq 0$. From these define six linear substitutions upon the coefficients $a_0, a_1, \cdots, a_n$, with $a_n \neq 0$, of the polynomial

$$A(z) = a_0 + a_1 z + a_2 z^2 + \cdots + a_n z^n$$

* Presented to the Society, San Francisco Section, April 9, 1921.

by means of the identities

$$\sigma A(z) \equiv \sigma'' A(\sigma' z) \equiv a_{\sigma 0} + a_{\sigma 1} z + a_{\sigma 2} z^2 + \cdots + a_{\sigma n} z^n ,$$

in which $A(\sigma' z)$ denotes the result of applying σ' to $A(z)$. Hence $\sigma A(z)$ is the polynomial derived from $A(z)$ by first operating with σ' and then multiplying the result by σ''. If $a_0 = 0$ the degree of $\sigma A(z)$ is $< n$ for some σ. The six σ's thus obtained are $1, \alpha, \beta, \gamma, \delta, \epsilon$, where α corresponds to α', α'', etc.

2. A set of generating relations for the cross ratio group is

$$\alpha'^3 = \gamma'^2 = (\alpha' \gamma')^2 = 1 .$$

We have $\alpha A(z) = (z-1)^n A(\alpha' z)$; whence

$$\alpha^2 A(z) \equiv \alpha(\alpha A(z)) = (z-1)^n [(z/(1-z))^n A(\alpha'^2 z)],$$

$$\alpha^3 A(z) \equiv \alpha(\alpha^2 A(z)) = (z-1)^n [(1/(z-1))^n A(\alpha'^3 z)] = A(z).$$

Hence $\alpha^3 = 1$, and similarly $\gamma^2 = 1$, $(\alpha\gamma)^2 = 1$, so that the σ form a group Γ simply isomorphic to the cross ratio group. For convenience of reference we reproduce its multiplication table, which is to be read in the usual way; thus $\alpha\gamma = \delta$, $\gamma\epsilon = \alpha$, and so on:

	1	α	β	γ	δ	ϵ
α	α	β	1	δ	ϵ	γ
β	β	1	α	ϵ	γ	δ
γ	γ	ϵ	δ	1	β	α
δ	δ	γ	ϵ	α	1	β
ϵ	ϵ	δ	γ	β	α	1

3. Obviously $a_{\gamma r} = a_{1n-r} \equiv a_{n-r}$. Let ρ, σ, τ be substitutions of Γ between which there is the relation $\rho\sigma = \tau$, and write $a_{\rho\sigma r}$ for the function of the coefficients a_r which is derived from the $a_{\sigma r}$ in the same way that $a_{\rho r}$ is from the a_r. Then $a_{\rho\sigma r} = a_{\tau r}$, and the coefficients of the six polynomials are related as shown by the table in § 2. If $\gamma\sigma = \tau$, then $a_{\tau r} = a_{\gamma\sigma r} = a_{\sigma n-r}$. Hence, writing as usual $0! = 1$, $\binom{r}{0} = \binom{0}{0} = 1$, $\binom{r}{s} = r!/(s!(r-s)!)$, we have the following forms of the $a_{\sigma r}$, $r = 0, 1, s, \cdots, n$, which explicitly define the substitutions of Γ:

$$a_{1r} = a_{\gamma n-r} = a_r ,$$

$$a_{\alpha r} = a_{\epsilon n-r} = (-1)^{n-r} \sum_{s=0}^{n-r} \binom{n-s}{r} a_s ,$$

$$a_{\beta r} = a_{\delta n-r} = (-1)^r \sum_{s=0}^{n-r} \binom{n-s}{r-s} a_{n-s} ,$$

the $a_{\sigma r}$ being given directly by the $\sigma A(z)$ for $\sigma = 1, \alpha, \beta$.

4. The set $A(z)$, $\alpha A(z)$, $\beta A(z)$, $\gamma A(z)$, $\delta A(z)$, $\epsilon A(z)$ is called anharmonic of degree n, n being the highest degree of any polynomial in the set. Any symmetric function of the polynomials $A(z)$, $\sigma A(z)$ is an invariant of the subgroup 1, σ ($\sigma = \gamma, \delta, \epsilon$); any symmetric function of $A(z)$, $\alpha A(z)$, $\beta A(z)$ is an invariant of the subgroup 1, α, β; and any symmetric function of $\sigma A(z)$ ($\sigma = 1, \alpha, \beta, \gamma, \delta, \epsilon$) is an invariant of Γ.

5. When $A(z)$ is reciprocal, $\gamma A(z) = A(z)$. Hence from the multiplication table $\delta A(z) = \alpha A(z)$, $\epsilon A(z) = \beta A(z)$, and in this case the anharmonic set reduces to $A(z)$, $\alpha A(z)$, $\beta A(z)$. We shall call such a set cyclic, and henceforth reserve the term anharmonic for sets that are not cyclic or, what is equivalent, contain no reciprocal polynomial. Any symmetric function of all the members of a cyclic set is invariant for 1, α, β.

6. The σ are determined by the coefficients a_r as in § 3 and are functions of the degree n of the set. When necessary to designate the polynomial fixing the σ we shall speak of the associated σ as the cross ratio substitution for that polynomial. To indicate that $A(z)$ belongs to a set of degree n we write $A_{(n)}(z)$, enclosing the n in ( ) to distinguish it from the rank defined in § 7. When it is a question of relations between the coefficients of polynomials of different degrees (§ 15) the a_r may be given double suffixes. Thus

$$A(z) \equiv A_{(n)}(z) \equiv \sum_{r=0}^{n} a_r z^r = \sum_{r=0}^{n} a_{nr} z^r$$

are merely different notations for one polynomial, and likewise for

$$\sigma A(z) \equiv \sigma A_{(n)}(z) \equiv \sum_{r=0}^{n} a_{\sigma r} z^r \equiv \sum_{r=0}^{n} a_{\sigma nr} z^r .$$

Where there can be no confusion we shall use the simpler forms.

7. Let x, z be independent variables, and put $\sqrt{z} = k$, $\sqrt{1 - z} = k'$, $i = \sqrt{-1}$. Let t, v be functions of x, k (or of x, z) such that

$$t(x, -k) = t(x, k) = -t(-x, k),$$

$$v(x, -k) = v(x, k) = v(-x, k),$$

and assume (cf. § 12) that these functions can be expanded in absolutely convergent power series in x of the form

$$t(x, k) = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n+1}}{(2n+1)!} T_{(n)}(z), \qquad v(x, k) = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n}}{(2n)!} V_{(n)}(z).$$

We now define $T_{(n)}(z)$ arising from an odd function generator $t(x, k)$ to be an odd polynomial of rank $2n + 1$, and $V_{(n)}(z)$ generated by an even function $v(x, k)$ to be an even polynomial of rank $2n$, and write

$$T_{(n)}(z) \equiv T_{2n+1}(z), \qquad V_{(n)}(z) \equiv V_{2n}(z),$$

so that the degree of a set containing a given polynomial is the greatest integer in half the rank. This change in notation is essential for the further development, as without it the application of the symbolic calculus of Blissard and Lucas is impracticable. Exhibiting the ranks rather than the degrees we have

$$t(x,k)=\sum_{n=0}^{\infty}\frac{(-1)^n x^{2n+1}}{(2n+1)!}T_{2n+1}(z)\equiv \sin T(z)x,$$

$$v(x,k)=\sum_{n=0}^{\infty}\frac{(-1)^n x^{2n}}{(2n)!}V_{2n}(z)\equiv \cos V(z)x,$$

the trigonometric forms being the purely symbolic equivalents of the series.

The principles of the symbolic method, which we shall use freely in the sequel, having been fully treated by Blissard, Lucas and Glaisher, in the works cited, we need not recall them here, except to emphasize the caution that in all operations with symbolic powers zero exponents must be included. Thus the first term of $\cos V(z)x$ is $V_0(z)$, not unity; the symbolic binomial $(p+q)^2=p^2q^0+2p^1q^1+p^0q^2\equiv p_2q_0+2p_1q_1+p_0q_2$, not $p_2+2p_1q_1+q_2$; and $(p-p)^2=2(p_2p_0-p_1^2)$, obtained from $(p-q)^2$ by putting $q=p$ in the final form of the latter.

8. That σ is one of the cross ratio substitutions for $A_n(z)$ may be indicated by writing σ_n, but for simplicity we shall put $\sigma_n A_n(z)\equiv \sigma A_n(z)$. Analogously to the sn, cn, ns, $\cdots$ notation for elliptic functions we denote each of the six polynomials $1A_n(z)\equiv A_n(z)$, $\alpha A_n(z)$, $\beta A_n(z)$, $\cdots$, $\epsilon A_n(z)$ by a double letter symbol $1A\equiv A$, αA, βA, $\cdots$, ϵA, and in any such σA regard the σ, A as inseparable. In Blissard's method $A_n(z)$ is written $A^n(z)$, or A^n when z is understood, the exponent being purely symbolic, and A is called an umbra. Similarly we now have $(\sigma A)^n\equiv \sigma A_n$, since σA is one symbol, and our umbræ are double-letter symbols σA. It is important to note once for all that $(\sigma A)^n$ is not $\sigma^n A_n$ in which σ^n has the usual meaning as a power of a substitution.

9. We require the operations transforming the generators of T, V into those of σT, σV respectively. Let $f(x,k)$ denote an arbitrary function of x, k, and Ω_i an operator which applied to f transforms it as follows:

$$\Omega_i f(x,k)=\phi_i(k)\,f(x\psi_i(k),\chi_i(k)).$$

We regard Ω_i as a tripartite operator,

$$\Omega_i\equiv\{\phi_i(k),\psi_i(k),\chi_i(k)\},$$

which replaces k in $f(x,k)$ by $\chi_i(k)$, x by $x\psi_i(k)$, and multiplies the function thus transformed by $\phi_i(k)$. Hence the product $\Omega_j\Omega_i$ in which Ω_i is applied first is

$$\{\phi_j(k)\phi_i(\chi_j(k)),\quad \psi_j(k)\psi_i(\chi_j(k)),\quad \chi_i(\chi_j(k))\};$$

and if $f(x, k)$ is odd in x, even in k,

$$\{\phi, \psi, \chi\} = \{-\phi, -\psi, \chi\} = \{\phi, \psi, -\chi\};$$

while if $f(x, k)$ is even in both x and k,

$$\{\phi, \psi, \chi\} = \{\phi, -\psi, \chi\} = \{\phi, \psi, -\chi\}.$$

Let σ', σ denote any cross ratio substitutions for $T_{2n+1}(z)$, $V_{2n}(z)$ respectively, so that symbolically

$$\sin \sigma' T(z)x \equiv \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n+1}}{(2n+1)!} \sigma' T_{2n+1}(z),$$

$$\cos \sigma V(z)x \equiv \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n}}{(2n)!} \sigma V_{2n}(z);$$

and designate by s', s operators such that

$$s' t(x, k) = \sin \sigma' T(z)x, \qquad sv(x, k) = \cos \sigma V(z)x.$$

Then by inspection of the series, the Greek and Latin letters (α, a), (α', a'), $\cdots$, (σ, s), (σ', s') corresponding, we have

$$\begin{aligned}
1 &= \{1, 1, k\}, & 1 &= \{1, 1, k\},\\
a' &= \{-i/k', ik', 1/k'\}, & a &= \{1, ik', 1/k'\},\\
b' &= \{-i/k, ik, ik'/k\}, & b &= \{1, ik, ik'/k\},\\
c' &= \{1/k, k, 1/k\}, & c &= \{1, k, 1/k\},\\
d' &= \{-i, i, k'\}, & d &= \{1, i, k'\},\\
e' &= \{1/k', k', ik/k'\}; & e &= \{1, k', ik/k'\}.
\end{aligned}$$

From the definitions the s' form a group G' which is simply isomorphic to Γ and whose multiplication table is obtained from that in § 2 upon replacing each Greek letter by its accented Latin correspondent; the multiplication table for the group G of operators s is obtained from G' by suppressing accents. The explicit forms of the operators given above can be verified easily on combining them according to the formulas developed for the Ω_i, noting that the s' refer to an odd, and the s to an even function.

10. Thus far s' has been applied only to the odd $t(x, k)$, and s to the even $v(x, k)$. The following cases are of equal importance, and they may be seen at once from the definitions. If

$$s' \equiv \{\phi'(k), \psi'(k), \chi'(k)\}, \qquad \text{then} \qquad s = \{1, \psi'(k), \chi'(k)\};$$

$$s' v(x, k) = \phi'(k) \cdot sv(x, k), \qquad st(x, k) = \frac{1}{\phi'(k)} \cdot s' t(x, k).$$

11. We now specialize the even V in each of two ways and the odd T in one way, getting in all for each integral degree $n \geqq 0$ a system of twelve polynomials distributed into two cyclic sets and one anharmonic. Indicating that the modulus of the elliptic functions is k by writing $\operatorname{sn}(x, k)$, etc., we define the fundamental polynomials S, C, P and their anharmonic transforms $\sigma'S$, σC, σP by their symbolic generators,

$$\operatorname{sn}(x, k) \equiv \sin S(z)x, \qquad s' \operatorname{sn}(x, k) \equiv \sin \sigma' S(z)x,$$

$$\operatorname{cn}(x, k) \equiv \cos C(z)x, \qquad s \operatorname{cn}(x, k) \equiv \cos \sigma C(z)x,$$

$$x \operatorname{ns}(x, k) \equiv \cos P(z)x; \qquad sx \operatorname{ns}(x, k) \equiv \cos \sigma P(z)x.$$

From the elements of elliptic functions the coefficients in $S_{2n+1}(z)$, $C_{2n}(z)$ are positive integers, $C_{2n}(z)$ is of degree $n-1$ in z, and S_{2n+1} is a reciprocal polynomial of degree n in z. Hence $P_{2n}(z)$ is a reciprocal polynomial, since $x \operatorname{ns} x = x/\operatorname{sn} x$. Therefore, omitting ranks from the notation, the coefficients in each of the sets $\sigma'S(z)$, $\sigma C(z)$ are integers, those in $\sigma P(z)$ are rational but not integral, and of the three sets $\sigma'S(z)$, $\sigma P(z)$ are cyclic, $\sigma C(z)$ is anharmonic.

12. If for all values of z whose absolute value does not exceed a constant different from zero the algebraic relation $F(z) = 0$ holds, then it is easily seen that $F(z') = 0$ is an identity in arbitrary z'. Hence in all polynomial formulas we consider z arbitrary, it being understood that when $F(z)$ is regarded as a coefficient in an infinite series z is such as to render the series absolutely convergent (all the series discussed have radius of convergence > 0), but that in all other connections z is arbitrary.

13. For convenience of comparison with Glaisher's grouping of the elliptic functions into triads we give the complete set of generators:

$$\operatorname{sn}(x, k) = \sin S(z)x, \qquad x \operatorname{ns}(x, k) = \cos P(z)x,$$

$$\operatorname{cn}(x, k) = \cos C(z)x, \qquad x \operatorname{cs}(x, k) = \cos \alpha P(z)x,$$

$$\operatorname{dn}(x, k) = \cos \gamma C(z)x; \qquad x \operatorname{ds}(x, k) = \cos \beta P(z)x;$$

$$\operatorname{dc}(x, k) = \cos \alpha C(z)x, \qquad \operatorname{cd}(x, k) = \cos \epsilon C(z)x,$$

$$\operatorname{nc}(x, k) = \cos \delta C(z)x, \qquad \operatorname{sd}(x, k) = \sin \beta S(z)x,$$

$$\operatorname{sc}(x, k) = \sin \alpha S(z)x; \qquad \operatorname{nd}(x, k) = \cos \beta C(z)x.$$

Glaisher points out that of the four triads the second is the most symmetrical, and that it should be taken as a basis for the jacobian elliptic functions rather than the traditional first. From the present point of view the second triad frequently appears to be wholly anomalous: while the sets $\sigma'S(z)$, $\sigma C(z)$ in many significant ways can be regarded as forming one complete system, the

set $\sigma P(z)$ stands apart in the majority of relations of a specific type. We see in a moment that $\sigma P(z)$ is related to the Bernoulli numbers, $\sigma' S(z)$, $\sigma C(z)$ to those of Genocchi and Euler respectively. Hence we have another instance of that superficial similarity and radical difference between the numbers of Bernoulli and Euler which has often been remarked.

14. When $z = 1$ and hence $k = 1$ the elliptic functions degenerate to circular functions of the gudermannian of x. From the values of these as given by Cayley, p. 59, and the symbolic generators for the B, E, G, R (numbers of Bernoulli, Euler, Genocchi and Lucas) in Lucas 5, p. 262, we find at once

$$\begin{aligned} x \operatorname{ns}(x,1) &= ix \cot ix = \cos 2Bix, \\ \operatorname{en}(x,1) &= \sec ix = \cos Eix, \\ \operatorname{dn}(x,1) &= \sec ix = \cos Eix, \\ 2x \operatorname{sn}(x,1) &= -2ix \tan ix = -\cos 2Gix, \\ x \operatorname{ds}(x,1) &= ix \operatorname{cosec} ix = 2\cos Rix, \\ x \operatorname{cs}(x,1) &= ix \operatorname{cosec} ix = 2\cos Rix. \end{aligned}$$

The even suffix notation of Lucas is used for all B, E, G, R, and the last two are defined by

$$G_{2n} = 2(1 - 2^{2n}) B_{2n}, \qquad R_{2n} = (1 - 2^{2n-1}) B_{2n}.$$

We require a fifth system of numbers H_{2n+1}, the so-called tangent coefficients, shown presently to be integers > 0,

$$H_{2n+1} = \frac{(-1)^{n+1}}{n+1} 2^{2n} G_{2n+2} = \frac{(-1)^n}{n+1} 2^{2n+1} (2^{2n+2} - 1) B_{2n+2}.$$

Comparing coefficients of like powers of x in the generators above with those in § 13 we have for $n \geqq 0$, $r > 0$ (note that in each case the degenerate form is expressed as a function of the rank $2n$ or $2n+1$),

$$\begin{aligned} &S_{2n+1}(1) = H_{2n+1}, && P_{2n}(1) = (-1)^n 2^{2n} B_{2n} = (2i)^{2n} B_{2n}, \\ &\alpha S_{2n+1}(1) = (-1)^n = -i \cdot i^{2n+1}, && \alpha P_{2n}(1) = 2(-1)^n R_{2n} = 2i^{2n} R_{2n}, \\ &\beta S_{2n+1}(1) = (-1)^n = -i \cdot i^{2n+1}; && \beta P_{2n}(1) = 2(-1)^n R_{2n} = 2i^{2n} R_{2n}; \\ &C_{2n}(1) = (-1)^n E_{2n} = i^{2n} E_{2n}, && \gamma C_{2n}(1) = (-1)^n E_{2n} = i^{2n} E_{2n}, \\ &\alpha C_0(1) = 1, \alpha C_{2r}(1) = 0, && \delta C_{2n}(1) = (-1)^n = i^{2n}, \\ &\beta C_{2n}(1) = (-1)^n = i^{2n}; && \epsilon C_0(1) = 1, \quad \epsilon C_{2r}(1) = 0. \end{aligned}$$

The coefficients in $S_{2n+1}(z)$, $C_{2n}(z)$ being integers > 0, so also are H_{2n+1}, $(-1)^n E_{2n}$.

15. An identity between some or all of the twelve elliptic functions implies and is implied by the identity between polynomials which is obtained upon equating coefficients of like powers of the argument x, and as degenerate cases for $z = 1$ of these identities we have relations between the numbers B, E, H, R. The elliptic identities may conveniently be segregated into classes according to the groups G, G' of § 9. All the identities arising from a given one by successive applications of the operations of G belong to the first class, all those similarly derived by means of G' belong to the second, and there are subsidiary classes corresponding to the cyclic subgroups of orders 3, 2. Polynomial identities derived from elliptic identities of a given class belong to one class. Elliptic identities are further subdivided into types according to the degree of the identity in sn, cn, $\cdots$, and the derived polynomial relations are similarly subdivided, the degenerate cases being included. By means of the table in § 2 and those next given, together with the formulas of § 10, we write down immediately from any elliptic identity all others of the same type, and hence on replacing the elliptic functions by their symbolic trigonometric equivalents from § 13 we at once infer all the polynomial relations of one type. It is clear that once the elliptic identity is given the rest of the process demands very little labor. In the next part we develop the symbolic method proper to the subject, and this still further reduces the algebra.

Consider any one of the polynomial relations. This is an identity in z and hence, z being arbitrary, it is equivalent to a set of identical relations between the coefficients of the several polynomials. The coefficients of $S_{2n+1}(z)$ are $n+1$ integers > 0 into which H_{2n+1} is partitioned; those of $C_{2n}(z)$ are n integers > 0 into which $(-1)^n E_{2n}$ is partitioned, and those of $P_{2n}(z)$ are a rational but not integral partition of $(-1)^n 2^{2n} B_{2n}$. From the explicit values of the transformed coefficients in § 3 the coefficients of the transforms of the fundamental polynomials S, C, P are known in terms of the foregoing partitions, and hence the polynomial relation is equivalent to a theorem concerning partitions of a certain kind of the numbers B, E, H of various ranks. Similar remarks apply to congruences between such of the polynomials as take integral values when z is an integer. This aspect is not further elaborated here, as its complete discussion presupposes a knowledge of the arithmetical form of the coefficients in $C_{2n}(z)$, $S_{2n+1}(z)$.* It is not difficult to give implicit arithmetical definitions of these coefficients, but this is not what is required.

* The well known method of Hermite (10, pp. 265, 269) for calculating $C_{2n}(z)$, $S_{2n+1}(z)$ does not give the required information, as recognized by Hermite himself (12, p. 237). His second solution (ibid.), as pointed out by the editors of his works, unfortunately is erroneous, and even if it were correct it is difficult to see what the general coefficient would be from the forms of those given. Writers on elliptic functions seem to have overlooked Hermite's remarks in the second citation, and to have assumed that his first paper is sufficient.

16. From the multiplication tables of G, G' combined with the results of the linear transformations of $\operatorname{sn}(x,k)$, $\operatorname{cn}(x,k)$ (as given for example by Glaisher (6, p. 120)) we have the following for the generators of the $\sigma' S$, σP:

	sn	sc	sd
1	sn	sc	sd
a'	sc	sd	sn
b'	sd	sn	sc

	x ns x	x cs x	x ds x
1	x ns x	x cs x	x ds x
a'	x cs x	x ds x	x ns x
b'	x ds x	x ns x	x cs x

the argument being x and the modulus k. Since each set is cyclic the transformations c', d', e' are respectively identical with 1, a', b'. Corresponding to § 5 we have: any symmetric function of all the members of either of these cyclic sets is an invariant of 1, a', b'.

For the anharmonic set the table is

	cn	dc	nd	dn	nc	cd
1	cn	dc	nd	dn	nc	cd
a	dc	nd	cn	nc	cd	dn
b	nd	cn	dc	cd	dn	nc
c	dn	cd	nc	cn	nd	dc
d	nc	dn	cd	dc	cn	nd
e	cd	nc	dn	nd	dc	cn

The modulus and argument in each case are k, x. Any symmetric function of the members of the following pairs (cn, dn), (dc, dn), (dn, nd) is an invariant of the group $(1, c)$, $(1, d)$ or $(1, e)$ respectively; any symmetric function of dn, nc, cd is an invariant of $(1, a, b)$, and any symmetric function of cn, dc, nd, dn, nc, cd is an invariant of G.

17. One example will suffice to show how all the relations of a given type may be written down from one of them. When k or k' occurs as a factor it is replaced in the final result by its z-equivalent. The modulus being k, consider $\operatorname{sn}^2 x + \operatorname{cn}^2 x = 1$. From this, since $\phi(k) = 1$ for each s,

$$(s \operatorname{sn} x)^2 + (s \operatorname{cn} x)^2 = 1.$$

Putting $s = a$ we have, by § 10 and the form of a' in § 9,

$$(ik' a' \operatorname{sn} x)^2 + (a \operatorname{cn} x)^2 = 1;$$

and hence from the tables in § 16, $-k'^2 \operatorname{sc}^2 x + \operatorname{dc}^2 x = 1$. To illustrate useful processes we shall also consider in detail the effects of operating with s'.

Let $s' \equiv \{\phi', \psi', \chi'\}$. Then

$$s'[\operatorname{sn}^2 x + \operatorname{cn}^2 x] = s' 1 = \phi'(k);$$

$$\phi'(k)[\operatorname{sn}^2(x\psi'(k), \chi'(k)) + \operatorname{cn}^2(x\psi'(k), \chi'(k))] = \phi'(k);$$

$$\operatorname{sn}^2(x\psi'(k), \chi'(k)) + \operatorname{cn}^2(x\psi'(k), \chi'(k)) = 1;$$

$$\frac{1}{\phi'^2(k)}(s' \operatorname{sn}(x, k))^2 + (s \operatorname{cn}(x, k))^2 = 1,$$

and hence for $s' = a'$, $-k'^2 \operatorname{sc}^2 x + \operatorname{dc}^2 x = 1$.

With a little practice the transform can be written down by inspection in any case directly from §§ 9, 10, 16. Thus the complete set here is

$$(1), \qquad \operatorname{cn}^2 x + \operatorname{sn}^2 x = 1,$$

$$(a \text{ or } a'), \qquad \operatorname{dc}^2 x - k'^2 \operatorname{sc}^2 x = 1,$$

$$(b \text{ or } b'), \qquad \operatorname{nd}^2 x - k^2 \operatorname{sd}^2 x = 1,$$

$$(c \text{ or } c'), \qquad \operatorname{dn}^2 x + k^2 \operatorname{sn}^2 x = 1,$$

$$(d \text{ or } d'), \qquad \operatorname{nc}^2 x - \operatorname{sc}^2 x = 1,$$

$$(e \text{ or } e'), \qquad \operatorname{cd}^2 x + k'^2 \operatorname{sd}^2 x = 1;$$

whence the final forms

$$\cos^2 C(z)x + \sin^2 S(z)x = 1,$$

$$\cos^2 \alpha C(z)x - (1 - z)\sin^2 \alpha S(z)x = 1,$$

$$\cos^2 \beta C(z)x - z \sin^2 \beta S(z)x = 1,$$

$$\cos^2 \gamma C(z)x + z \sin^2 S(z)x = 1,$$

$$\cos^2 \delta C'(z)x - \sin^2 \alpha S(z)x = 1,$$

$$\cos^2 \epsilon C'(z)x + (1 - z)\sin^2 \beta S(z)x = 1.$$

Simple rules may be devised for writing down the appropriate multipliers such as z, $1 - z$ above, from the forms of the s', s and § 10, for the several terms of any symbolic identity when the argument z of the polynomials is transformed by the substitutions of Γ. As these present no difficulty we omit them.

II. Isomorphism with elliptic functions

18. To find the polynomial relation equivalent to a given elliptic identity we evidently must consider the properties of products of t symbolic sines and r symbolic cosines in the cases $r, t > 0$; $r > 0$, $t = 0$; $r = 0$, $t > 0$. When several symbolic factors of a product are identical we proceed as in the following example. Let λ, μ denote umbræ (§ 8), and suppose the coefficient of x^n in $\cos^2 \lambda x$ is required. We write $\cos^2 \lambda x \equiv \cos \lambda x \cos \mu x$, find the coefficient of x^n in $\cos \lambda x \cos \mu x$ by actual multiplication of the series for $\cos \lambda x$, $\cos \mu x$ or otherwise (§ 24), and in the result, *after* each exponent of λ, μ has been degraded

to a suffix, replace μ by λ. Once more we emphasize that in all expansions zero exponents must be included. Thus $\lambda, \mu, \cdots, \nu$ being umbræ,

$$(\lambda + \mu + \cdots + \nu)^n = \sum \frac{n!}{r!\,p!\cdots t!}\lambda^r \mu^p \cdots \nu^t = \sum \frac{n!}{r!\,p!\cdots t!}\lambda_r \mu_p \cdots \nu_t,$$

the $\sum$ extending to all $r \geqq 0, p \geqq 0, \cdots, t \geqq 0$ such that

$$r + p + \cdots + t = n.$$

As always, $0! = 1$, and after the completion of all formal operations (multiplications, divisions, etc.) $\lambda^r \mu^p \cdots \nu^t$ is to be replaced by $\lambda_r \mu_p \cdots \nu_t$. By convention, for *any* particular choice of the signs,

$$(\pm \lambda^l \pm \mu^m \pm \cdots \pm \nu^n)^0 = \lambda^0 \mu^0 \cdots \nu^0 \equiv \lambda_0 \mu_0 \cdots \nu_0.$$

19. All letters $\lambda, \cdots, \pi'$ denote umbræ, and the sets $\Lambda, \Lambda', Z, Z' \equiv (\lambda, \mu, \cdots, \rho), (\lambda', \mu', \cdots, \tau'), (\zeta, \xi, \cdots, \phi), (\zeta', \xi', \cdots, \pi')$ contain respectively r, t, f, p letters. For $n \geqq 0$ write

$$(\Lambda|\Lambda')^n \equiv \sum \pm (\pm \lambda \pm \mu \pm \cdots \pm \rho \pm \lambda' \pm \mu' \pm \cdots \pm \tau')^n,$$

in which all the exponents are symbolic, the summation extends to the 2^{r+t} possible combinations of signs within the parentheses, and the outer sign in each case is the product of the signs of $\lambda', \mu', \cdots, \tau'$. The important special cases $t = 0$, $r = 0$ give respectively

$$(\Lambda|\,)^n = \sum (\pm \lambda \pm \mu \pm \cdots \pm \rho)^n,$$
$$(\,|\Lambda')^n = \sum \pm (\pm \lambda' \pm \mu' \pm \cdots \pm \tau')^n.$$

The *umbral factors* of $(\Lambda|\Lambda')^n$ are by definition $\lambda, \mu, \cdots, \rho, \lambda', \mu', \cdots, \tau'$, and similarly for the others. If in these the letters be interpreted as ordinary quantities and the exponents as algebraic we have precisely the symmetric functions considered by Kronecker, cf. § 26.

By definition the respective types of $(\Lambda|\Lambda')^n$, $(\Lambda|\,)^n$, $(\,|\Lambda')^n$ are $(r|t)$, $(r|0)$, $(0|t)$, and the weight of each is n, = the sum of the suffixes in the final non-symbolic forms of each. Considered as functions of $\lambda_j, \cdots, \tau'_k$ these final forms are homogeneous of degrees $r + t$, r, t respectively, the degree in any case being equal to the total number of letters in the bar function $(\Lambda|\Lambda')^n$, etc. The properties of these functions $(\Lambda|\Lambda')^n, \cdots$ are immediate from the expressions in § 20 for their symbolic generators. From the principles of the symbolic calculus as developed by Blissard and Lucas it is evident that products, etc., of symbolic sines and cosines can be combined formally according to the rules of trigonometry, and that the coefficients of like powers of x in the several transforms thus effected of any identity are equal. By starting from the elliptic functions of a pure imaginary argument the development can be

carried out isomorphically to the theory of the hyperbolic instead of the circular functions. This in some respects is preferable, but being committed to the other by Lucas' trigonometric generators for B, E, G, R in § 14 we shall not follow it.

20. Replacing each symbolic sine or cosine by its exponential equivalent we see immediately that

$$2^{r+t}\cos\lambda x\cos\mu x\cdots\cos\rho x\sin\lambda' x\sin\mu' x\cdots\sin\tau' x$$

$$=(-1)^{\frac{t-1}{2}}\sin(\Lambda|\Lambda')x \quad\text{or}\quad (-1)^{\frac{t}{2}}\cos(\Lambda|\Lambda')x$$

according as t is odd or even;

$$2^{r}\cos\lambda x\cos\mu x\cdots\cos\rho x=\cos(\Lambda|\,)x;$$

and according as t is odd or even,

$$2^{t}\sin\lambda' x\sin\mu' x\cdots\sin\tau' x=(-1)^{\frac{t-1}{2}}\sin(\,|\Lambda')x \quad\text{or}\quad (-1)^{\frac{t}{2}}\cos(\Lambda'|\,)x,$$

all of which are included in the first. Hence $(\Lambda|\Lambda')^n$ is an even function of each of the r letters in Λ, and an odd function of each of the t letters in Λ'; or, $(\Lambda|\Lambda')^n$ of type $(r|t)$ is r-fold even and t-fold odd. Moreover $(\Lambda|\Lambda')^n$ is symmetric in the letters in Λ, also in those in Λ'. Again, from the generators, $(\Lambda|\Lambda')^{2n}=0$ if t is odd, while $(\Lambda|\Lambda')^{2n+1}=0$ if t is even. The corresponding statements for $(\Lambda|\,)^n$, $(|\Lambda')^n$ are included as special cases.

21. In $(\Lambda|\Lambda')^n$ we are concerned with (1) the umbræ $\lambda,\cdots,\tau'$; (2) the type $(r|t)$; (3) the weight n. Each of these has a species of addition theorem. Denote by (Λ, Z) the set consisting of all the letters in Λ together with all those in Z, and similarly for (Λ', Z'). Then $((\Lambda, Z)|(\Lambda', Z'))^n$ is of type $((r+f)|(t+p))$, and is symmetric in Λ, Z, also in Λ', Z'. The addition theorems with respect to types are given by the following, all of which are obvious on remarking that the generator of $((\Lambda, Z)|(\Lambda', Z'))^n$ is the product of the generators of $(\Lambda|\Lambda')^n$, $(Z|Z')^n$, and reapplying the several cases of the generators in § 20 to these products before equating coefficients of like powers of x. Let $(n, t, p)\equiv(n', t', p')\bmod 2$; then

(n', t', p')	$4((\Lambda, Z)\|(\Lambda', Z'))^n$	$((r+f)\|(t+p))$
$(0, 0, 0)$	$((\Lambda\|\Lambda'), (Z\|Z')\|)^n$	$((r\|t), (f\|p)\|)$
$(0, 1, 1)$	$(\|(\Lambda\|\Lambda'), (Z\|Z'))^n$	$(\|(r\|t), (f\|p))$
$(1, 0, 1)$	$((\Lambda\|\Lambda')\|(Z\|Z'))^n$	$((r\|t)\|(f\|p))$
$(1, 1, 0)$	$((Z\|Z')\|(\Lambda\|\Lambda'))^n$	$((f\|p)\|(r\|t))$

For example, when n, t, p are all even, a function of type $((r+f)|(t+p))$ is expressible linearly in terms of functions of types $(r|t)$, $(f|p)$ in the manner

shown in the second column. The calculation of such a function is thus reduced to a series of symbolic multiplications (and subsequent additions) of functions whose types are $(r|t)$, $(f|p)$.

22. Omitting the addition theorems for the umbræ we restrict the discussion of those for the weights to the particularly important cases of two and three letters. The general case of n umbræ is treated similarly. In all that follows multiplications indicated by dots are purely symbolic. These are performed analogously to algebraic multiplications but by the addition of suffixes instead of exponents. Thus

$$(\lambda_r + \mu_r)\cdot(\lambda_t + \mu_t) = \lambda_r\cdot\lambda_t + \lambda_r\cdot\mu_t + \mu_r\cdot\lambda_t + \mu_r\cdot\mu_t$$
$$= \lambda_{r+t} + \lambda_r\,\mu_t + \mu_r\,\lambda_t + \mu_{r+t};$$

$\lambda_2\cdot\lambda_3 = \lambda_5$; $\lambda_2\cdot\lambda_2 = \lambda_4$, not λ_2^2. Algebraic multiplication is a special case of this.

Taking first the case of two umbræ λ, μ we define auxiliary functions $\{\}$ by

$$2\{\lambda, \mu|\}^r \equiv \phi^r + \psi^r, \qquad 2\{|\lambda, \mu\}^r \equiv \phi^r - \psi^r,$$
$$\phi \equiv \lambda + \mu, \qquad \psi \equiv \lambda - \mu, \qquad r \geqq 0,$$

and put $L^r \equiv L_r = \{\lambda, \mu|\}^r$, $M^r \equiv M_r = \{|\lambda, \mu\}^r$. Then

$$4L^{2r} = (\lambda, \mu|)^{2r}, \qquad 4M^{2r} = (|\lambda, \mu)^{2r}, \qquad 4M^{2r+1} = (\lambda|\mu)^{2r+1}.$$

We have $\phi^r = L^r + M^r$, $\psi^r = L^r - M^r$, and hence for $t \geqq 0$,

$$\phi^{r+t} = (L_r + M_r)\cdot(L_t + M_t), \qquad \psi^{r+t} = (L_r - M_r)\cdot(L_t - M_t).$$

On the other hand

$$2L^{r+t} = \phi^{r+t} + \psi^{r+t}, \qquad 2M^{r+t} = \phi^{r+t} - \psi^{r+t}.$$

Whence, substituting for ϕ^{r+t}, ψ^{r+t} and degrading exponents, we have the addition theorems

$$L_{r+t} = L_r\cdot L_t + M_r\cdot M_t, \qquad M_{r+t} = L_r\cdot M_t + M_r\cdot L_t.$$

For example $L_1 = \lambda_1\,\mu_0$, $L_2 = \lambda_2\,\mu_0 + \mu_2\,\lambda_0$, $M_1 = \lambda_0\,\mu_1$, $M_2 = 2\lambda_1\,\mu_1$; hence $L_1\cdot L_2 = \lambda_3\,\mu_0 + \lambda_1\,\mu_2$, $M_1\cdot M_2 = 2\lambda_1\,\mu_2$. Again

$$L_3 = \tfrac{1}{2}[(\lambda + \mu)^3 + (\lambda - \mu)^3] = \lambda_3\,\mu_0 + 3\lambda_1\,\mu_2,$$

so that $L_3 = L_1\cdot L_2 + M_1\cdot M_2$ as required by the first of the theorems.

23. The case* of three umbræ λ, μ, ν is treated in the same way, and we need

* If λ, μ, ν be interpreted as ordinaries and exponents as algebraic, the special cases of the addition theorems of the first kind (in § 21) for the functions $\{\lambda, \mu, \nu|\}^r$, etc., which are of the same form as those for $(\lambda, \mu, \nu|)^r$, etc., have interesting consequences when r is prime for Fermat's quotients. Cf. Bachmann, Journal für Mathematik, vol. 142 (1913), pp. 41–50; Dickson, *History of the Theory of Numbers*, vol. I, p. 111.

give only the results. There can be no confusion between $\alpha, \beta, \gamma, \delta$ here and the substitutions of Γ, nor between B and the Bernoulli numbers. Write

$$\alpha = \lambda + \mu + \nu, \qquad 4A^r \equiv 4\{\lambda, \mu, \nu|\}^r \equiv \alpha^r + \beta^r + \gamma^r + \delta^r,$$
$$\beta = \lambda - \mu + \nu, \qquad 4B^r \equiv 4\{\lambda, \mu|\nu\}^r \equiv \alpha^r + \beta^r - \gamma^r - \delta^r,$$
$$\gamma = \lambda + \mu - \nu, \qquad 4C^r \equiv 4\{|\lambda, \mu, \nu\}^r \equiv \alpha^r - \beta^r - \gamma^r + \delta^r,$$
$$\delta = \lambda - \mu - \nu, \qquad 4D^r \equiv 4\{\nu|\lambda, \mu\}^r \equiv \alpha^r - \beta^r + \gamma^r - \delta^r,$$

and put $A^r, B^r, C^r, D^r \equiv A_r, B_r, C_r, D_r$. We have

$$SA^{2r} = (\lambda, \mu, \nu|)^{2r}, \qquad SB^{2r+1} = (\lambda, \mu|\nu)^{2r+1},$$
$$SC^{2r+1} = (|\lambda, \mu, \nu)^{2r+1}, \quad SD^{2r} = (\nu|\lambda, \mu)^{2r},$$

and the addition theorem for $r, t \geqq 0$,

$$A_{r+t} = A_r \cdot A_t + B_r \cdot B_t + C_r \cdot C_t + D_r \cdot D_t,$$
$$B_{r+t} = A_r \cdot B_t + B_r \cdot A_t + C_r \cdot D_t + D_r \cdot C_t,$$
$$C_{r+t} = A_r \cdot C_t + B_r \cdot D_t + C_r \cdot A_t + D_r \cdot B_t,$$
$$D_{r+t} = A_r \cdot D_t + B_r \cdot C_t + C_r \cdot B_t + D_r \cdot A_t.$$

There also are multiplication theorems for the calculation of A_{nr}, etc., but we shall omit these.

24. As they are frequently useful we write down the $\{\}$ equivalents for 2 and 3 letters of the generators in § 20:

$$\cos \lambda x \cos \mu x = \cos \{\lambda, \mu|\}x = \cos \{\mu, \lambda|\}x,$$
$$\sin \lambda x \sin \mu x = -\cos \{|\lambda, \mu\}x = -\cos \{|\mu, \lambda\}x,$$
$$\cos \lambda x \sin \mu x = \sin \{\lambda|\mu\}x = \sin \{\mu, \lambda|\}x,$$
$$\sin \lambda x \cos \mu x = \sin \{\mu|\lambda\}x = \sin \{\lambda, \mu|\}x;$$
$$\cos \lambda x \cos \mu x \cos \nu x = \cos \{\lambda, \mu, \nu|\}x = \cos \{\{\lambda, \mu|\}, \nu|\}x,$$
$$\cos \lambda x \cos \mu x \sin \nu x = \sin \{\lambda, \mu|\nu\}x = \sin \{\nu, \{\lambda, \mu|\}|\}x,$$
$$\sin \lambda x \sin \mu x \sin \nu x = -\sin \{|\lambda, \mu, \nu\}x = -\sin \{\nu, \{|\lambda, \mu\}|\}x,$$
$$\sin \lambda x \sin \mu x \cos \nu x = -\cos \{\nu|\lambda, \mu\}x = -\cos \{\{|\lambda, \mu\}, \nu|\}x.$$

25. None of the twelve polynomials can be computed by linear recurrence, that is, by an ordinary difference equation of finite order, since otherwise an elliptic function would satisfy a linear differential equation of finite order and the first degree. We now give two symbolic linear recurrences by means of which, in conjunction with the foregoing addition theorems, the actual computation of all the polynomials can be effected systematically without reference to elliptic functions. Or, by using the results of §§ 30, 31, the polynomials can

be calculated directly from the recurrences. The addition theorems in any case shorten the labor but are not indispensable.

To find the recurrences for L, M of § 22 we proceed as if λ, μ were ordinaries, forming the equation in w whose roots are ϕ, ψ and multiplying the result by w^n, finally degrading exponents and replacing ordinary products by dot multiplications (§ 22). The formal algebraic details parallel those for the case of ordinaries, as given for example by Lucas 5, pp. 308–310, and may be omitted. For $w =$ either L or M, and $n \geqq 0$, we find

$$w_{n+2} - 2\lambda_1 \cdot w_{n+1} + (\lambda_2 - \mu_2) \cdot w_n = 0,$$

with the initial values directly from the definitions of L, M,

$$L_0 = \lambda_0 \mu_0, \qquad L_1 = \lambda_1 \mu_0, \qquad M_0 = 0, \qquad M_1 = \lambda_0 \mu_1.$$

Thus for $n = 0$, $w = L$, we have

$$\begin{aligned} L_2 = 2\lambda_1 \cdot L_1 - (\lambda_2 - \mu_2) \cdot L_0 &= 2\lambda_1 \cdot \lambda_1 \mu_0 - (\lambda_2 - \mu_2) \cdot \lambda_0 \mu_0 \\ &= 2\lambda_2 \mu_0 - \lambda_2 \mu_0 + \lambda_0 \mu_2 = \lambda_2 \mu_0 + \lambda_0 \mu_2, \end{aligned}$$

which is correct. Put $n = 1$:

$$L_3 = 2\lambda_1 \cdot (\lambda_2 \mu_0 + \lambda_0 \mu_2) - (\lambda_2 - \mu_2) \cdot \lambda_1 \mu_0 = \lambda_3 \mu_0 + 3\lambda_1 \mu_2,$$

and so on.

Similarly if w is any one of A, B, C, D,

$$\begin{aligned} w_{n+4} - 4\lambda_1 \cdot w_{n+3} + 2(3\lambda_2 - \mu_2 - \nu_2) \cdot w_{n+2} - 4(\lambda_3 - \lambda_1 \mu_2 - \mu_2 \nu_1) \cdot w_{n+1} \\ + (\lambda_4 + \mu_4 + \nu_4 - 2\lambda_2 \mu_2 - 2\mu_2 \nu_2 - 2\nu_2 \lambda_2) \cdot w_n = 0; \end{aligned}$$

$$A_0 = \lambda_0 \mu_0 \nu_0, \qquad B_0 = 0,$$

$$A_1 = \lambda_1 \mu_0 \nu_0, \qquad B_1 = \lambda_0 \mu_0 \nu_1,$$

$$A_2 = \lambda_2 \mu_0 \nu_0 + \lambda_0 \mu_2 \nu_0 + \lambda_0 \mu_0 \nu_2, \qquad B_2 = 2\lambda_1 \mu_0 \nu_1,$$

$$A_3 = \lambda_3 \mu_0 \nu_0 + 3\lambda_1 \mu_2 \nu_0 + 3\lambda_1 \mu_0 \nu_2; \qquad B_3 = \lambda_0 \mu_0 \nu_3 + 3\lambda_2 \mu_0 \nu_1 + 3\lambda_0 \mu_2 \nu_1;$$

$$C_0 = 0, \qquad D_0 = 0,$$

$$C_1 = 0, \qquad D_1 = \lambda_0 \mu_1 \nu_0,$$

$$C_2 = 2\lambda_0 \mu_1 \nu_1, \qquad D_2 = 2\lambda_1 \mu_1 \nu_0,$$

$$C_3 = 6\lambda_1 \mu_1 \nu_1; \qquad D_3 = \lambda_0 \mu_3 \nu_0 + 3\lambda_2 \mu_1 \nu_0 + 3\lambda_0 \mu_1 \nu_2.$$

Provided multiplications be interpreted as above, viz., as dot multiplications, the theory of such symbolic recurrences is obviously identical in its formal aspects with that of ordinary recurrences. Hence for L, M we have at once isomorphs of all the algebraic formulas developed by Lucas for his U, V. Thus, for example, L, M are expressible as symbolic continuants, cf. Lucas 4,

p. 193. Similarly A, B, C, D have symbolic properties completely analogous to the algebraic relations between any system of four independent solutions of

$$y_{n+4} + py_{n+3} + qy_{n+2} + ry_{n+1} + sy_n = 0;$$

in particular there are symbolic equivalents of the generalized continued fractions of Jacobi, Fürstenau, and others. This can be continued in the same way for functions of $4, 5, \cdots, n$ umbræ.

There is a more general aspect of the formulas of this part which is useful elsewhere. The development of any function of x which is even, odd, or arbitrary in x can be written in the symbolic forms

$$\cos \eta x, \quad \sin \zeta x, \quad \cos \eta x + i \sin \zeta x$$

respectively, and the coefficient of x^n in the product of any number of such functions can be most readily investigated by the symbolic trigonometry which we have sketched. When n functions are concerned the symbolic recurrences are of order n.

One special case of interest may be mentioned. As remarked in § 19, $(\Lambda\,|\,)^n$, $(|\Lambda')^n$, when exponents are interpreted as algebraic and the umbræ as ordinaries, are the symmetric functions of Kronecker 13, p. 385, which he took as his point of departure for deducing certain properties of the Bernoulli numbers. The step-by-step symbolic interpretation of Kronecker's analysis may be made without difficulty, and all his formulas translated into terms of umbræ instead of ordinaries. The formulas thus derived contain his as limiting cases. Bernoulli numbers enter Kronecker's formulas as coefficients in the expansion of hyperbolic tangents, and hence the manner in which they appear is quite distinct from that of the present discussion.

26. Corresponding to any product of elliptic functions there is a unique product of symbolic sines and cosines, and hence a unique $(\Lambda\,|\,\Lambda')^n$. From any elliptic identity we write down an identity between bar functions of general odd or even weight. When the identity contains ns x, cs x or ds x it is first multiplied throughout by the lowest power of x such that these functions can be replaced by x ns x, x cs x or x ds x respectively wherever they occur. Now the umbræ $\lambda, \mu, \cdots, \tau'$ in $(\Lambda\,|\,\Lambda')^n$ are its umbral factors (§ 19), and hence the isomorphism between elliptic and polynomial identities is complete.

III. Algebraic relations and congruences

27. For $z = 1$ the polynomials degenerate to Bernoulli and allied numbers, § 14. Let ω denote a complex cube root of unity. The cases $z = -\omega$, $z = \frac{1}{2}$, which by an obvious analogy may be called the equianharmonic and harmonic, are noted here in passing. They are characterized by the vanishing

of certain invariants (§ 4). Write

$$A_n(z),\ \alpha A_n(z),\ \beta A_n(z) \equiv A_n(z),\ A_n'(z),\ A_n''(z),$$

$$I_n(z) = \begin{vmatrix} A_n(z) & A_n'(z) & A_n''(z) \\ A_n'(z) & A_n''(z) & A_n(z) \\ A_n''(z) & A_n(z) & A_n'(z) \end{vmatrix}.$$

Then from the table in § 2,

$$I_n(z) = \begin{vmatrix} A_n(z) & A_n'(z) & A_n''(z) \\ \alpha A_n(z) & \alpha A_n'(z) & \alpha A_n''(z) \\ \beta A_n(z) & \beta A_n'(z) & \beta A_n''(z) \end{vmatrix},$$

and from the values of α', β' in § 1, it is evident that this vanishes when $z = \alpha' z$, $z = \beta' z$, $\alpha' z = \beta' z$, that is, when $z = -\omega$, $-\omega^2$. Hence

$$I_n(-\omega) = I_n(-\omega^2) = 0,$$

and therefore from the tables in §§ 13, 16 we have, when $u = -\omega$ or $-\omega^2$,

$$S^3_{2n+1}(u) + \alpha S^3_{2n+1}(u) + \beta S^3_{2n+1}(u) - 3S_{2n+1}(u)\,\alpha S_{2n+1}(u)\,\beta S_{2n+1}(u) = 0,$$

$$P^3_{2n+1}(u) + \alpha P^3_{2n+1}(u) + \beta P^3_{2n+1}(u) - 3P_{2n+1}(u)\,\alpha P_{2n+1}(u)\,\beta P_{2n+1}(u) = 0,$$

$$C^3_{2n}(u) + \alpha C^3_{2n}(u) + \beta C^3_{2n}(u) - 3C_{2n}(u)\,\alpha C_{2n}(u)\,\beta C_{2n}(u) = 0,$$

$$\gamma C^3_{2n}(u) + \epsilon C^3_{2n}(u) + \delta C^3_{2n}(u) - 3\gamma C_{2n}(u)\,\epsilon C_{2n}(u)\,\delta C_{2n}(u) = 0,$$

the fourth of which follows from the third on transforming by γ, or it is independently obvious from the last table in § 16. The product of the last two is another invariant of the same form which vanishes for the same values of u. This invariant corresponds to the entire last table in § 16, the determinant of which is a circulant of the sixth order, and hence, by a well known theorem of Glaisher, is expressible as a circulant of the third order. From the values of α'', β'' in § 1 and the form of $I_n(z)$ each of these invariants vanishes also when $z = 0, 1$. The harmonic invariants are written down in the same way.

28. Applying the first formulas of § 24 to the final set in § 17 we have, for $n > 0$,

$$\{C, C|\}^{2n} - \{|S, S\}^{2n} = 0,$$

$$\{\alpha C, \alpha C|\}^2 + (1-z)\{|\alpha S, \alpha S\}^{2n} = 0,$$

$$\{\beta C, \beta C|\}^{2n} + z\{|\beta S, \beta S\}^{2n} = 0,$$

$$\{\gamma C, \gamma C|\}^{2n} - z\{|S, S\}^{2n} = 0,$$

$$\{\delta C, \delta C|\}^{2n} + \{|\alpha S, \alpha S\}^{2n} = 0,$$

$$\{\epsilon C, \epsilon C|\}^{2n} - (1-z)\{|\beta S, \beta S\}^{2n} = 0,$$

in which $C = C(z)$, and likewise for the rest. For $z = 1$ we get from § 14 the corresponding degenerate form, ($i = \sqrt{-1}$),

$$\{iE, iE|\}^{2n} - \{|H, H\}^{2n} = 0, \quad \text{or} \quad (-1)^n\{E, E|\}^{2n} = \{|H, H\}^{2n},$$

from the first or fourth of the above. The remaining degenerate forms are identities between powers of i. In full the degenerate relation is

$$(-1)^n \sum_{r=0}^{n} \binom{2r}{2r} E_{2n-2r} E_{2r} = \sum_{r=1}^{n} \binom{2n}{2r-1} H_{2n-2r+1} H_{2r-1}.$$

Similarly for the polynomials P with $n > 1$,

$$\{P, P|\}^{2n} - \{\beta P, \beta P|\}^{2n} = 0, \qquad \{P, P|\}^{2n} - \{\alpha P, \alpha P|\}^{2n} = 0,$$
$$2^{2n-2}\{B, B|\}^{2n} - \{R, R|\}^{2n} = 0,$$

the first of which corresponds to

$$x^2 \operatorname{ns}^2 x - x^2 \operatorname{ds}^2 x = x^2 k^2,$$

and the second comes from this as shown in § 17 by transforming with respect to α. The degenerate relation reduces by § 14 to

$$\sum_{r=0}^{n} \binom{2r}{2r} (2^{2n-2r-1} + 2^{2r-1} - 1) B_{2n-2r} B_{2r} = 0.$$

29. It is well known that all the algebra including the addition theorems of the elliptic functions follows from the square relations $\operatorname{cn}^2 x + \operatorname{sn}^2 x = 1$, etc., and the expressions for the derivatives of $\operatorname{sn} x$, $\operatorname{cn} x$. Hence all relations between the polynomials are implicit in § 28 and the formulas next given, which are written down from the expressions for the derivatives of $\operatorname{sn} x$, $\operatorname{cn} x$, x ns x and their transforms by § 16, or directly from the first member of each of the three sets as suggested at the end of § 17. The first formulas of § 24 are used for all, and the results may be checked at a glance by comparing with Glaisher 6, p. 92. Thus, the dot being as in § 22, we obtain from

$$\operatorname{sn}(x, k) = \sin S(z)x,$$

by differentiation,

$$\operatorname{cn}(x, k) \operatorname{dn}(x, k) = S_1(z) \cdot \cos S(z)x,$$

or

$$\cos C(z)x \cos \gamma C(z)x = S_1(z) \cdot \cos S(z)x,$$

which gives at once the first of the following:

$$S_{2n+1} = \{C, \gamma C|\}^{2n}, \qquad (2n+1)P_{2n+2} = -\{\alpha P, \beta P|\}^{2n+2},$$
$$\alpha S_{2n+1} = \{\alpha C, \delta C|\}^{2n}, \qquad (2n+1)\alpha P_{2n+2} = -\{P, \beta P|\}^{2n+2},$$
$$\beta S_{2n+1} = \{\beta C, \epsilon C|\}^{2n}; \qquad (2n+1)\beta P_{2n+2} = -\{P, \alpha P|\}^{2n+2};$$
$$C_{2n+2} = \{S, \gamma C|\}^{2n+1}, \qquad \gamma C_{2n+2} = z\{S, C|\}^{2n+1},$$
$$\alpha C_{2n+2} = -(1-z)\{\alpha S, \delta C|\}^{2n+1}, \qquad \delta C_{2n+2} = -\{\alpha S, \alpha C|\}^{2n+1},$$
$$\beta C_{2n+2} = -z\{\beta S, \epsilon C|\}^{2n+1}; \qquad \epsilon C_{2n+2} = (1-z)\{\beta S, \beta C|\}^{2n+1},$$

in all of which $n \geqq 0$. By means of the L, M addition theorems (§ 22) and recurrences (§ 25) all of the polynomials can be calculated successively, and without excessive labor, from the initial values for $n = 0, 1$ (which are given by the definitions). Putting $z = 1$ as before we find the degenerate forms:

$$H_{2n+1} = (-1)^n\{E, E|\}^{2n}, \qquad (2n+1)\,2^{2n}\,B_{2n+2} = -\{R, R|\}^{2n+2},$$
$$E_{2n+2} = (-1)^{n+1}\{H, iE|\}^{2n}, \qquad (2n+1)\,R_{2n+2} = -\{2B, R|\}^{2n+2};$$

whence, by comparison with those in § 29,

$$H_{2n+1} = \{|H, H\}^{2n}, \qquad (2n+1)\,B_{2n+2} = -\{B, B|\}^{2n+2}.$$

30. As a last example of relations involving bar functions of not more than two umbræ we take the identities which express the elliptic functions as products of two others, and those between the functions and their reciprocals. The argument is z, as before.

$$S_{2n+1} = \{\alpha S, C|\}^{2n+1} = \{\beta S, \gamma C|\}^{2n+1},$$
$$\alpha S_{2n+1} = \{\beta S, \alpha C|\}^{2n+1} = \{S, \delta C|\}^{2n+1},$$
$$\beta S_{2n+1} = \{S, \beta C|\}^{2n+1} = \{\alpha S, \epsilon C|\}^{2n+1};$$
$$P_{2n} = \{\alpha P, \delta C|\}^{2n} = \{\beta P, \beta C|\}^{2n},$$
$$\alpha P_{2n} = \{\beta P, \epsilon C|\}^{2n} = \{P, C|\}^{2n},$$
$$\beta P_{2n} = \{P, \gamma C|\}^{2n} = \{\alpha P, \alpha C|\}^{2n};$$

$$(2n+1)\,C_{2n} = \{S, \alpha P|\}^{2n+1}, \qquad C_{2n} = \{\gamma C, \epsilon C|\}^{2n},$$
$$(2n+1)\,\alpha C_{2n} = \{\alpha S, \beta P|\}^{2n+1}, \qquad \alpha C_{2n} = \{\delta C, \gamma C|\}^{2n},$$
$$(2n+1)\,\beta C_{2n} = \{\beta S, P|\}^{2n+1}, \qquad \beta C_{2n} = \{\epsilon C, \delta C|\}^{2n},$$
$$(2n+1)\,\gamma C_{2n} = \{S, \beta P|\}^{2n+1}, \qquad \gamma C_{2n} = \{C, \alpha C|\}^{2n},$$
$$(2n+1)\,\delta C_{2n} = \{\alpha S, P|\}^{2n+1}, \qquad \delta C_{2n} = \{\alpha C, \beta C|\}^{2n},$$
$$(2n+1)\,\epsilon C_{2n} = \{\beta S, \alpha P|\}^{2n+1}, \qquad \epsilon C_{2n} = \{\beta C, C|\}^{2n};$$

whence the degenerate forms,

$$H_{2n+1} = (-1)^n\{1, E|\}^{2n+1}, \qquad \{1, E|\}^{2n} = 0, \quad n > 0,$$
$$1 = (-1)^n\{H, i|\}^{2n+1}, \qquad (2n+1)\,E_{2n} = 2(-1)^n\{H, iR|\}^{2n+1},$$
$$2^{2n-1}\,B_{2n} = \{R, 1|\}^{2n}, \qquad \{1, R|\}^{2n+1} = 0, \quad n > 0,$$
$$2n+1 = \{1, 2B|\}^{2n+1}, \qquad 2R_{2n} = \{2B, E|\}^{2n}.$$

31. The simplest general relations between polynomials of one kind and of several ranks are found from the differential equations for the elliptic functions. Let λ, μ, ν denote umbræ and consider functions v, t, w, u of x such that

$$v = \cos \lambda x, \qquad v'' = av + 2bv^3,$$
$$t = \sin \mu x, \qquad t'' = ct + 2dt^3,$$
$$w = xu = \cos \nu x, \qquad u'' = gu + 2hu^3,$$

in which a, b, c, d, g, h are independent of x, and double accents denote second derivatives with respect to x. Then v, t, u are elliptic functions, and hence λ, μ, ν polynomials in z (= the square of the modulus). From the differential equations for v, t, w we get by substituting the symbolic trigonometric equivalents of v, t, w, performing the differentiations and equating coefficients of like powers of x,

$$\lambda_{2n+2} + a\lambda_{2n} + 2b\{\lambda, \lambda, \lambda|\}^{2n} = 0,$$
$$\mu_{2n+3} + c\mu_{2n+1} - 2d\{|\mu, \mu, \mu\}^{2n+1} = 0,$$
$$n(2n+1)\nu_{2n+2} + (n+1)(2n+1)g\nu_{2n} - h\{\nu, \nu, \nu|\}^{2n+2} = 0,$$

in which we have used the last formulas of § 24 for the coefficients* in v^3, t^3, w^3. Knowing the values of the constants in the cases $\lambda = C, \mu = S, \nu = P$ from Glaisher 6, p. 122, we write down the rest of the following table by the methods of § 17. The argument is z throughout.

$(\lambda, a, b) =$	$(\mu, c, d) =$	$(\nu, g, h) =$
$(C, 2z-1, -z)$,	$(S, -1-z, z)$,	$(P, -1-z, 1)$,
$(\alpha C, -1-z, 1)$,	$(\alpha S, 2-z, 1-z)$,	$(\alpha P, 2-z, 1)$,
$(C, 2-z, -1+z)$,	$(\beta S, 2z-1, -z+z^2)$;	$(\beta P, 2z-1, 1)$.
$(\gamma C, 2-z, -1)$,		
$(\delta C, 2z-1, 1-z)$,		
$(\epsilon C, -1-z, z)$;		

We omit the degenerate forms. The bar functions of λ, μ, ν may be computed by the addition theorems of § 23 and the recurrences of § 25, and hence each of the polynomials for all ranks can be calculated independently of the remaining polynomials. In using §§ 23, 25 we perform all operations for three distinct umbræ, putting these equal to one another in the successive final steps (see § 18). We must pass on to a brief discussion of the congruences for prime moduli.

* The differential equation for w is $x^2 w'' = 2xw' + (gx^2 - 2)w + hw^3$.

32. Henceforth p denotes an odd prime > 0. (Some of the congruences are also valid modulo 2, but this case is of such slight interest that we ignore it.) The theorem of arithmetic which gives the congruences modulo p is due to Lucas 4, pp. 229–230:

$$\binom{m}{n} \equiv \binom{m_1}{n_1}\binom{m_1'}{n_1'} \qquad \text{mod } p,$$

$$m = m_1 p + m_1', \qquad n = n_1 p + n_1', \qquad 0 \leqq m_1', \quad n_1' < p.$$

We shall in future omit the "mod p" in writing congruences, and use "$\equiv$" only in the sense of congruence modulo p. From Lucas' theorem it is easy to infer that, λ, μ denoting umbræ of integers and the significance of the dot being as in § 22,

$$(\lambda \pm \mu)^{rp+s} \equiv (\lambda^p \pm \mu^p)^r \cdot (\lambda \pm \mu)^s, \qquad 0 \leqq s < p,$$

the upper signs or the lower being taken throughout. By repeated application of this we find the following. Let

$$N = r_n p^n + r_{n-1} p^{n-1} + r_{n-2} p^{n-2} + \cdots + r_1 p_1 + r_0$$

be the (unique) expression of N in the scale of p, so that $0 \leqq r_j < p$ ($j = 0, 1, \cdots, n$). Then

$$(\lambda + \mu)^N \equiv (\lambda^{p^n} + \mu^{p^n})^{r_n} \cdot (\lambda^{p^{n-1}} + \mu^{p^{n-1}})^{r_{n-1}} \cdots (\lambda^p + \mu^p)^{r_1} \cdot (\lambda + \mu)^{r_0},$$

$$(\lambda - \mu)^N \equiv (\lambda^{p^n} - \mu^{p^n})^{r_n} \cdot (\lambda^{p^{n-1}} - \mu^{p^{n-1}})^{r_{n-1}} \cdots (\lambda^p - \mu^p)^{r_1} \cdot (\lambda - \mu)^{r_0}.$$

Denote the right-hand members of these congruences by

$$\prod_{j=0}^{n}{}' (\lambda^{p^{n-j}} \pm \mu^{p^{n-j}})^{r_{n-j}}$$

respectively, in which the accented Π' indicates that the products of the several factors are to be performed as dot multiplications. Then

$$\{\lambda, \mu|\}^N \equiv \tfrac{1}{2}\prod_{j=0}^{n}{}' (\lambda^{p^{n-j}} + \mu^{p^{n-j}})^{r_{n-j}} + \tfrac{1}{2}\prod_{j=0}^{n}{}' (\lambda^{p^{n-j}} - \mu^{p^{n-j}})^{r_{n-j}},$$

$$\{|\lambda, \mu\}^N \equiv \tfrac{1}{2}\prod_{j=0}^{n}{}' (\lambda^{p^{n-j}} + \mu^{p^{n-j}})^{r_{n-j}} - \tfrac{1}{2}\prod_{j=0}^{n}{}' (\lambda^{p^{n-j}} - \mu^{p^{n-j}})^{r_{n-j}}.$$

The distributed *non-symbolic* equivalent of each right-hand member after the performance of all dot multiplications and the degradation of all exponents is a quadratic form with positive integral coefficients of the form

$$\binom{r_n}{r_n'}\binom{r_{n-1}}{r_{n-1}'} \cdots \binom{r_0}{r_0'} \qquad (r_n' \leqq r_n, \quad r_{n-1}' \leqq r_{n-1}, \quad \cdots, \quad r_0' \leqq r_0),$$

and the term of which this is the coefficient is

$$\lambda_{N-N'}\,\mu_{N'} \qquad (N' = r'_n\,p^n + r'_{n-1}p^{n-1} + \cdots + r'_0)\,.$$

That the coefficients are positive is evident from the left-hand members in which all coefficients are such; the coefficients on the right are the positive residues modulo p of those on the left.

If $\rho_N = \{\lambda, \mu|\}^N$, or if $\rho_N = \{|\lambda, \mu\}^N$, the above congruences are applicable in an obvious manner, equalities of this sort having been considered in the preceding sections. From § 15 it is clear that any congruence between polynomials (the P set does not immediately enter the discussion since the coefficients are not integers) is equivalent to a system of congruences between the linear functions of the form in § 3 into which H, E and their anharmonic transforms (the degenerate forms of the corresponding polynomials) for different ranks are partitioned.

Two important special cases of the general formula in this section are

$$\{\lambda, \mu|\}^p \equiv \lambda_p\,\mu_0, \qquad \{|\lambda, \mu\}^p \equiv \lambda_0\,\mu_p\,.$$

Hence in all cases we have the residue of a bar function of two integral umbræ when the rank of the function is equal to or greater than the modulus. The case of two umbræ will be completed for a prime modulus when we find, as next, the residue when the rank is less than the modulus.

33. Restating another theorem of Lucas 4, p. 229, in a form adapted to our purpose we have

$$\binom{p-1}{n} \equiv (-1)^n, \qquad \binom{p-r}{n} \equiv (-1)^n \binom{n+r-1}{r-1}$$

$$(0 < r < p, \quad n < p - r)\,.$$

Hence for $r > 0$,

$$(\lambda + \mu)^{p-r} \equiv \sum_{s=0}^{p-r} (-1)^s \binom{r+s-1}{s} \lambda_{p-r-s}\,\mu_s\,,$$

$$(\lambda - \mu)^{p-r} \equiv \sum_{s=0}^{p-r} \binom{r+s-1}{s} \lambda_{p-r-s}\,\mu_s;$$

and hence, including all cases (cf. § 24), for $h \geqq 0$,

$$\{\lambda, \mu|\}^{p-2h-1} \equiv \sum_{s=0}^{(p-2h-1)/2} \binom{2h+2s}{2s} \lambda_{p-2h-2s-1}\,\mu_{2s}\,,$$

$$\{\lambda, \mu|\}^{p-2h-2} \equiv \sum_{s=0}^{(p-2h-3)/2} \binom{2h+2s+1}{2s} \lambda_{p-2h-2s-2}\,\mu_{2s}\,,$$

$$\{\lambda, \mu\}^{p-2h-1} \equiv -\sum_{s=0}^{(p-2h-1)/2} \binom{2h+2s-1}{2s-1} \lambda_{p-2h-2s}\, \mu_{2s-1}.$$

34. The congruences in § 33 admit of immediate generalization to any number m of umbræ $\lambda, \rho, \cdots, \mu$. It is easily seen in the same way, or as a consequence of the fundamental formula of § 33,

$$(\lambda + \mu)^{rp+s} \equiv (\lambda^p + \mu^p)^r \cdot (\lambda + \mu)^s,$$

that we have

$$(\lambda + \rho + \cdots + \mu)^{rp+s} \equiv (\lambda^p + \rho^p + \cdots + \mu^p)^r \cdot (\lambda + \rho + \cdots + \mu)^s;$$

and hence in the previous notation,

$$(\lambda + \rho + \cdots + \mu)^N \equiv \prod_{j=0}^{n}{}' (\lambda^{p^{n-j}} + \rho^{p^{n-j}} + \cdots + \mu^{p^{n-j}})^{r_{n-j}},$$

with $2^m - 1$ similar congruences in which the signs of some or all of $\lambda, \rho, \cdots, \mu$ are changed throughout. From these the congruences for the general bar functions of § 19 follow at once. The particular case $m = 3$ is of interest in connection with the formulas of § 32, but the length of this paper precludes further discussion. From the general formulas given there is no difficulty in writing down the congruences for bar functions of two or three letters from the algebraic relations developed earlier in the paper. For $z = 1$ some of the degenerate cases are well known, and hence provide checks. The other degenerate cases may be checked directly in a similar manner.

References

1. J. Blissard: *Theory of generic equations*, Quarterly Journal of Mathematics, vol. 4 (1861), pp. 279–305; vol. 5 (1862), pp. 58–75, 185–208.
2. J. Blissard: *Examples of the use and application of the representative notation*, ibid., vol. 6 (1864), pp. 49–64.
3. J. Blissard: *Researches in analysis*, ibid., vol. 7 (1865), pp. 155–170, 223–226.
4. E. Lucas: *Théorie des fonctions numériques simplement périodiques*, American Journal of Mathematics, vol. 1 (1878), pp. 184–238, 289–321.
5. E. Lucas: *Théorie des Nombres*, 1891.
6. J. W. L. Glaisher: *On elliptic functions*, Messenger of Mathematics, vol. 11 (1881–1882), pp. 81–95, 120–138.
7. J. W. L. Glaisher: *On the Bernoullian function*, Quarterly Journal of Mathematics, vol. 29 (1898), pp. 1–168 (especially pp. 122–168).
8. J. W. L. Glaisher: *On a class of relations connecting any n consecutive Bernoullian functions*, ibid., vol. 42 (1911), pp. 86–157.
9. A. Cayley: *An Elementary Treatise on Elliptic Functions*, Edition 2 (1895).
10. C. Hermite: *Remarques sur le développement de* cos am x, Paris Comptes Rendus, vol. 67 (1863), p. 613; Journal de Mathématiques pures et appliquées, (2) vol. 9 (1864), p. 289; Œuvres, vol. 2, pp. 265–270.
11. C. Hermite: *Sur les développements de* $F(x) = \mathrm{sn}^a\, x\, \mathrm{cn}^b\, x\, \mathrm{dn}^c\, x$ *où les exposants sont entiers*, Académie royale des Sciences de Stockholm, Bihang III, No. 10 (1875), pp. 3–10; Œuvres, vol. 3, pp. 222–235.

12. C. Hermite: *Sur le développement des fonctions elliptiques*, Journal für die reine und angewandte Mathematik, vol. 81 (1876), pp. 220–228; Œuvres, vol. 3, pp. 236–245.
13. L. Kronecker: *Sur quelques fonctions symétriques et sur les nombres de Bernoulli*, Journal de Mathématiques pures et appliquées, (2) vol. 1 (1856), pp. 385–391.
14. L. Kronecker: *Ueber die Bernoullischen Zahlen*, Journal für die reine und angewandte Mathematik, vol. 94 (1883), pp. 268–269.

University of Washington,
Seattle, Wash.

NEW PROPERTIES OF ALL REAL FUNCTIONS*

BY

HENRY BLUMBERG

INTRODUCTION

In a former paper† the author communicated a number of properties of every real function $f(x)$; these were stated in terms of the successive saltus functions associated with a given function. The present paper makes no use of the saltus functions, and the new properties are direct qualifications of $f(x)$. Since $f(x)$ is entirely unrestricted,—except, of course, that it is defined‡ and therefore finite for every real x—these qualifications are consequences of nothing else than that $f(x)$ is a function. A new light is thus thrown upon the nature of a function.

The new properties are of two types, descriptive and metric; the former are concerned with density, and the latter with measure (Lebesgue).

For the sake of greater concreteness of exposition, we shall discuss, for the most part, planar sets and real functions of two real variables.

1. DESCRIPTIVE PROPERTIES

We shall say that a planar set S is an *I-region* (= open set) if every point of S is an inner point of S; i.e., no point of S is the limit of a sequence of points not in S.

We shall deal with binary relations $\mathfrak{R}$ between I-regions and points. $I\mathfrak{R}P$ shall mean that the I-region I has the relation $\mathfrak{R}$ to the point P. The relation $\mathfrak{R}$ is said to be *closed*, if the relationships $I\mathfrak{R}P_n$ and $\lim_{n\to\infty} P_n = P$ imply $I\mathfrak{R}P$. By a *neighborhood* of a point P, we understand an I-region containing P; by a *partial neighborhood* of P, an I-region of which P is an inner or boundary point. A neighborhood of P is, therefore, also a partial neighborhood of P. We have the following

LEMMA I. *If $\mathfrak{R}$ is a closed relation, then the points for which (a) $N\mathfrak{R}P$ for every neighborhood N of P, and (b) a partial neighborhood $N_<$ exists such that $N_<\,\mathfrak{R}P$ (i.e., $N_<\,\mathfrak{R}P$ is false) constitute a non-dense (i.e., nowhere dense) set M.*

* Presented to the Society March 29, 1919, and December 30, 1920.

† *Certain general properties of functions*, Annals of Mathematics, vol. 18 (1917), p. 147.

‡ Even this restriction may be partially dispensed with; cf. Section 2.

Proof: Since $\mathfrak{R}$ is closed and $N_< \mathfrak{R}P$, it follows that no sequence of points P_n exists such that $\lim P_n = P$ and $N_< \mathfrak{R}P_n$ for every n. Hence a neighborhood N of P exists such that $N_< \mathfrak{R}A$ for every point A of N. In particular $N_< \mathfrak{R}B$ for every point B of $NN_<$, the set of points common to N and $N_<$. Since $N_<$ is a neighborhood of B, it follows that no B belongs to M. Every neighborhood of a point P of M thus contains a subset, which is an I-region, every point of which is outside of M.

Now let $z = f(x, y)$ be any given real function of the two real variables x and y, defined for every point of the plane Π, which we take to be the XY plane of a cartesian system of coördinates. We shall write also $z = f(P)$, where $P = (x, y)$ ranges over Π. For every pair of real numbers (x, y), there is a single number $f(x, y)$; otherwise f is unrestricted. We define as follows the relation $\mathfrak{R}_{r_1 r_2}$, where r_1 and r_2 are two real numbers and $r_1 < r_2$: *If P is a point and I an I-region of* Π, *then* $I\mathfrak{R}_{r_1 r_2} P$, *if and only if an infinite sequence of points P_n of I exists, such that*

$$\lim P_n = P, \qquad \lim f(P_n) \text{ exists},$$

and

$$r_1 \leqq \lim f(P_n) \leqq r_2.$$

It follows that $\mathfrak{R}_{r_1 r_2}$ is closed. For suppose that $\lim P^{(n)} = P$ and $I\mathfrak{R}_{r_1 r_2} P^{(n)}$ for every n. Then, for every n, there exists a sequence

$$\{P_m^{(n)}\}, \qquad m = 1, 2, \cdots,$$

of points of I such that $\lim_{m \to \infty} P_m^{(n)} = P^{(n)}$ and $r_1 \leqq \lim_{m \to \infty} f(P_m^{(n)}) \leqq r_2$. From the $P_m^{(n)}$, we may, in view of the preceding relations, select a sequence of points $\{Q_n\}$, $n = 1, 2, \cdots$, such that

$$\lim Q_n = P \qquad \text{and} \qquad r_1 \leqq \liminf f(Q_n) \leqq \limsup f(Q_n) \leqq r_2;$$

and from the sequence $\{Q_n\}$ a subsequence $\{P_n\}$ such that $\lim f(P_n)$ exists. We then have

$$\lim P_n = P \qquad \text{and} \qquad r_1 \leqq \lim f(P_n) \leqq r_2.$$

$\mathfrak{R}_{r_1 r_2}$ is therefore closed.

By the use of Lemma I, we may thus conclude that for every function $f(x, y)$ and every pair of numbers $r_1 < r_2$, the points P of the XY plane for which (*a*) $N\mathfrak{R}_{r_1 r_2} P$ for every neighborhood N of P and (*b*) $\overline{N_< \mathfrak{R}_{r_1 r_2} P}$ for some partial neighborhood $N_<$ of P constitute a non-dense set $T_{r_1 r_2}$. Let T be the sum of all the sets $T_{r_1 r_2}$, r_1 and r_2 ($r_1 < r_2$) ranging independently over all the rational numbers. T, being the sum of a denumerable number of non-dense sets, is exhaustible* (i.e., of first category according to Baire).

* For the terminology cf. Denjoy, Journal de Mathématiques, ser. 7, vol. 1 (1915), pp. 122–125.

We shall now use this property of T to obtain a property of all functions. For this purpose, we introduce the notion of *dense approach*. The function $f(x, y)$ is said to be *densely approached at the point* (ξ, η), or in other words, the point (ξ, η, ζ), $\zeta = f(\xi, \eta)$, of the "surface" $z = f(x, y)$ is said to be densely approached, if for every positive ϵ there exists a planar neighborhood N of (ξ, η), such that the points (x, y) of N for which $|f(x, y) - f(\xi, \eta)| < \epsilon$ form a dense set in N.

We have the following

THEOREM I. *For every real function $f(x, y)$ whatsoever, the points of the surface $z = f(x, y)$ that are densely approached form a residual set (= complement of an exhaustible set). Conversely, given any residual set R whatsoever, a function $f(x, y)$ exists that is densely approached at and only at the points of R.*

*Proof.** The points of Π'—i.e., of the surface $z = f(x, y)$—are either isolated or not. The isolated points of Π' form a denumerable set; therefore the points of Π whose corresponding points of Π' are isolated form a denumerable and therefore an exhaustible set. It is consequently sufficient to prove that the points P of Π, for which P' is a limit point of Π' but is not densely approached by Π', form an exhaustible set. Since P' is not densely approached, there exists an ϵ such that for every neighborhood N of P, the points Q of N for which $|f(Q) - f(P)| < \epsilon$ are not dense in N; i.e., an I-region exists in every neighborhood N of P such that $|f(Q) - f(P)| \geqq \epsilon$ for every point Q of the I-region. Hence there exists a partial neighborhood $N_<$ of P such that the inequality $|f(P) - f(Q)| \geqq \epsilon$ holds for every point Q of $N_<$. Let now r_1 be a rational number between $f(P) - \epsilon$ and $f(P)$; and r_2 a rational number between $f(P)$ and $f(P) + \epsilon$. Then it follows from the inequality $|f(Q) - f(P)| \geqq \epsilon$ for all points Q of $N_<$ that the set $\{Q'\}_{Q \text{ in } N_<}$ —i.e., the set of points of Π' corresponding to the points Q of $N_<$—has no point (ξ, η, z) as a limit with $r_1 \leqq z \leqq r_2$, where $(\xi, \eta) = P$. Hence $N_< \, \mathfrak{R}_{r_1 r_2} P$. On the other hand, P' is a limit point of Π' and hence of every N' where N is a neighborhood of P; hence N' has (ξ, η, ζ), with $r_1 < \zeta = f(\xi, \eta) < r_2$, as a limit point, and P is therefore a point of the non-dense set $T_{r_1 r_2}$ associated with the closed relation $\mathfrak{R}_{r_1 r_2}$, and hence of the exhaustible set T which is the sum of all the $T_{r_1 r_2}$, r_1 and r_2 ranging independently over all the rational numbers. The points of Π' that are not densely approached thus constitute an exhaustible set.

The proof of the converse is immediate. For let E be the exhaustible set complementary to the given residual set R. We may write

$$E = E_1 + E_2 + \cdots E_n + \cdots$$

* If S is any subset of Π, we shall understand by S' the set of points of the surface $z = f(x, y)$ corresponding to the points of S. Thus Π' represents the totality of surface points. Analogously, P' will represent the surface point corresponding to P.

where the E_n's are non-dense in II and no pair of E_n's have common points. Let

$$f(P) = \frac{1}{n}, \qquad \text{if} \qquad P \text{ is in } E_n,$$

and

$$f(P) = 2, \qquad \text{if} \qquad P \text{ is in } R.$$

It is then clear that $f(P)$ is densely approached, if and only if P is in R.

Theorem I shows what a remarkable degree of "regularity" every function possesses. This may be better realized, perhaps, by using the following equivalent definition of dense approach. The function $f(x, y)$ is said to be densely approached at P, if for every partial neighborhood $N_<$ of P, $N'_<$ has P' as a limit. We thus get the following equivalent theorem—we omit a restatement of the converse—which shows a kind of "microscopic symmetry" in the structure of the surface $z = f(x, y)$ for unconditioned f.

Theorem I'. *With every function $f(x, y)$ whatsoever, there is associated a residual set R—dependent on f—of the XY plane such that if $P = (\xi, \eta)$ is a point of R, and $N_<$ a partial neighborhood of P, the set $N'_<$, which consists of the points of the surface $z = f(x, y)$ that correspond to the points of $N_<$, has $(\xi, \eta, f(\xi, \eta))$ as a limit point.*

We pass now to further discriminations in the manner of "approach." The function f is said to be *inexhaustibly approached* at the point P of II, or in other words, P' is inexhaustibly approached by II', if every neighborhood of P contains, for every $\epsilon > 0$, an inexhaustible set of points—i.e., a set that is not exhaustible,—at which f differs from $f(P)$ by less than ϵ.

The function f is said to be *exhaustibly approached*—we then say also that the point $P' = (x, y, f(x, y))$ is exhaustibly approached by II'—at the point $P = (x, y)$, if it is not inexhaustibly approached at the point; in other words, if a neighborhood N of (x, y) and a number $\epsilon > 0$ exist, such that the points of N where f differs from $f(x, y)$ by less than ϵ form an exhaustible set. If M is any planar set, we shall use, in connection with approach, the expression "via M" to indicate that (x, y) is restricted to range in M. Thus "f is inexhaustibly approached at P via M" means that for every neighborhood N of P and every $\epsilon > 0$, the set MN, which is the aggregate of points common to M and to N, contains an inexhaustible set of points at which f differs from $f(P)$ by less than ϵ. The following definition for exhaustible approach is equivalent to the one above: The point P' of II' is exhaustibly approached, if a sphere S exists with P' as center such that the points of II corresponding to the points of II'S form an exhaustible set.

We have the following

Theorem II. *For every function $f(x, y)$, there exists in the XY plane a*

residual set R, dependent on f, such that if P is a point of R, and $N_<$ a partial neighborhood of P, the function f is inexhaustibly, and therefore densely, approached at P via $RN_<$.*

Proof. We first note that *the points of* II *where f is exhaustibly approached form an exhaustible set E_1.* For space contains a denumerable dense set

$$A = \{A_1, A_2, \cdots, A_m, \cdots\}.$$

Associate with each A_m a sequence of spheres S_{mn}, $n = 1, 2, \cdots$, with A_m as center and $1/n$ as radius. Blacken every sphere S_{mn}—the interior as well as the boundary—which is such that the totality of points of II′ contained in it corresponds to an exhaustible subset of II. Let E_1 be the subset of II constituted by the points which correspond to points of II′ that lie in one or more black spheres. Since the number of black spheres is at most denumerable and each black sphere contributes an exhaustible set to E_1, it follows that E_1 is exhaustible. If P is a point at which f is exhaustibly approached, P' is in at least one black sphere; hence P belongs to E_1.

Now suppose we remove from II the set E_1 of points at which f is exhaustibly approached, thus obtaining the set $R_1 = \text{II} - E_1$ of points where f is inexhaustibly approached. Since the values of f in an exhaustible set cannot affect the property of inexhaustible approach, every point of R_1 is inexhaustibly approached via R_1. Furthermore the points of R_1, at which f is densely approached via R_1, form a residual set of II. For assume for the present that f is bounded,—we shall later drop this restriction. Let k be any number greater than the least upper bound of f. Let the function $g(x, y)$ be equal to k at the points of E_1 and equal to $f(x, y)$ at the points of R_1. Since the set of points at which $g(x, y)$ is densely approached forms a residual set R_g of which only an exhaustible subset can lie in E_1, there must be a residual set of II in R_1 at the points of which g is densely approached. This dense approach is clearly valid via R_1 and for the function f. Let E_2 be the exhaustible subset of points of R_1 at which f is not densely approached via R_1. It will now be seen that at every point of the set $R = \text{II} - E_1 - E_2$, f is densely and inexhaustibly approached via R, and furthermore that f is inexhaustibly approached at every point P of R via $N_< R$, the common part of $N_<$ and R, where $N_<$ is any partial neighborhood of P. Every point of R is inexhaustibly approached via R_1 and hence via R, since an exhaustible set has no effect upon inexhaustible approach. Moreover every point P of R is densely approached via R_1; i.e., if $N_<$ is a partial neighborhood of P, and ϵ a positive number, there exists a point Q of $N_< R_1$ at which f differs from $f(P)$ by less than ϵ. Since, however, f is inexhaustibly approached at Q via R, and since R

* f is densely approached at P via M, which implicitly has P as limit, if for every partial neighborhood $N_<$ of P such that $N_< M$ has P as limit, the set $(N_< M)'$, i.e., the set of surface points corresponding to the points of $N_< M$, has P' as a limit.

differs from R_1 by an exhaustible set, it follows that Q' is a limit point of R', and therefore N_ϵ contains a point of R at which f differs from $f(P)$ by less than ϵ. P is thus densely and inexhaustibly approached via $N_\epsilon R$.

If f is unbounded, we use the transformation

$$\bar{f}(x, y) = \frac{f(x, y)}{1 + |f(x, y)|}.$$

The new function $\bar{f}$ is bounded, but the properties of dense and of exhaustible and of inexhaustible approach are preserved by this transformation. We may therefore drop the restriction made above in reference to boundedness.

THEOREM III. *With every function $f(x, y)$ there is associated (not uniquely, however) a dense set D of the XY plane such that $f(x, y)$ is continuous, if (x, y) ranges over D.*

Proof. We shall say that a set of points is "*ϵ-spaced*" (in the plane II) if every circle of radius ϵ contains at least one point of the set. Let $\epsilon_1, \epsilon_2, \epsilon_3, \cdots$ be a decreasing sequence of positive numbers with $\lim \epsilon_m = 0$. Let K_1 be any set that is everywhere dense in II and at the points of which $f(x, y)$ is densely approached via K_1; such a set exists according to Theorem II. Since K_1 is everywhere dense in II, we may select in K_1 a subset

$$D_1 = \{P_{11}, P_{12}, \cdots, P_{1n}, \cdots\},$$

which is isolated (i.e., no point of D is a limit point of D) and ϵ_1-spaced. We enclose each P_{1n} in a circle C_{1n} in such a way that (*a*) no two C_{1n}'s overlap—for this purpose it is sufficient to make the radius of C_{1n} less than one half the greatest lower bound of the distances from P_{1n} to the other points of D_1; and that (*b*) C_{1n} contains a dense subset of K_1 at every point of which the value of f differs from $f(P_{1n})$ by less than ϵ_1—the existence of such a subset for a sufficiently small radius of C_{1n} is guaranteed by the fact that f is densely approached at P_{1n} via K_1. Denote by J_{1n} the set of all the points in K_1 that are in the interior of C_{1n} and at which f differs from $f(P_{1n})$ by less than ϵ_1, and by K_2 the subset of K_1 constituted by all the points of all the J_{1n} and the points of K_1 that lie in the interior of $\mathrm{II} - \sum_{n=1}^{\infty} C_{1n}$. At every point Q of K_2 f is densely approached via K_2. For if Q is an interior point of $\mathrm{II} - \sum C_{1n}$, this property is obvious. If, however, Q is in C_{1n}, let $|f(Q) - f(P_{1n})| = \epsilon_1 - \delta$, $\delta > 0$. f is densely approached at Q via points of K_1 at which f differs from $f(Q)$ by less than δ; that is, via points where f differs from $f(P_{1n})$ by less than ϵ_1; hence via K_2. Since K_2 is everywhere dense and f is densely approached via K_2 at every one of its points, we may treat it as we did K_1. We select an isolated, ϵ_2-spaced subset $D_2 = \{P_{21}, P_{22}, \cdots, P_{2n}, \cdots\}$ of K_2 such that D_2 contains every point of D_1. We now enclose each P_{2n} in a circle C_{2n} in such a way that (*a*) C_{2n} contains a dense subset of points of K_2

where the value of f differs from $f(P_{2n})$ by less than ϵ_2; (*b*) no two C_{2n} overlap; and (*c*) C_{2n} lies either entirely within a C_{1n} or entirely outside of all C_{1n}. Denote by J_{2n} the set of all the points of C_{2n} K_2 at which f differs from $f(P_{2n})$ by less than ϵ_2, and by K_3 the subset of K_2 constituted by all the points of all the J_{2n} and the points of K_2, lying neither in the interior nor on the boundary of any C_{2n}. The set K_3, like K_2, is everywhere dense and such that at every point of it f is densely approached via K_3, and the process may therefore be continued. The set

$$D_m = \{P_{m1}, P_{m2}, \cdots, P_{mn}, \cdots\}$$

is an ϵ_m-spaced, isolated subset of K_m and a superset of D_{m-1}. The circle C'_{mn} enclosing P_{mn} (*a*) contains a dense subset of points of K_m where f differs from $f(P_{mn})$ by less than ϵ_m, (*b*) has no points in common with C'_{mp} for $p \neq n$ and (*c*) lies entirely outside of all C_{pq}, $p < m$, if P_{mn} is outside of all C_{pq} for $p < m$, and entirely inside every C_{pq}, $p < m$, containing P_{mn}.

We now define the set D of the theorem to be the sum of the D_m's ($m = 1, 2, \cdots$). That D is dense follows from the fact that D_m is ϵ_m-spaced. If P_{mn} is a point of D, it lies in every circle of a sequence of circles C_{mn}, $C_{m+1,\,n^{(1)}}, \cdots$, such that at the points of $C_{m+p,\,n^{(p)}}$ that belong to D the function f differs from $f(P_{mn})$ by less than ϵ_{m+p}. This proves the asserted continuity.

2. Generalizations

So far we have dealt only with real functions of two real variables. It is apparent that the considerations apply equally well to functions of a single variable and to functions of n real variables. The results are, however, essentially of a still more general character. Without entering upon extreme refinements of generalization we may note that the arguments in the preceding section are substantially valid for every function $f(P)$, defined in a set $\mathfrak{S}$ satisfying the following conditions:

(1) $\mathfrak{S}$ is a metric;* that is to say, with every pair of elements P and Q of $\mathfrak{S}$ there is associated a non-negative, real number $\overline{PQ}$ (Fréchet's *écart*) in such a way that if P, Q, and R are any three elements of $\mathfrak{S}$, then

(*a*) $\overline{PQ} = \overline{QP}$;

(*b*) $\overline{PQ} = 0$, when and only when $P = Q$; and

(*c*) $\overline{PQ} + \overline{QR} \geqq \overline{PQ}$.

(2) $\mathfrak{S}$ is a complete space (vollständiger Raum);† that is to say, if $\{P_1, P_2,$

* Cf., for example, Fréchet, *Sur quelques points du calcul fonctionnel*, Rendiconti del Circolo Matematico di Palermo, vol. 22 (1907), p. 1, and Hausdorff, *Grundzüge der Mengenlehre*, 1914, p. 211.

† Hausdorff, loc. cit., p. 315.

$\cdots, P_n, \cdots\}$ is a "regular" sequence of elements of $\mathfrak{S}$, in other words, if for every $\epsilon > 0$ there exists an integer n_ϵ such that $\overline{P_\lambda P_\mu} < \epsilon$ for $\lambda > n_\epsilon$ and $\mu > n_\epsilon$, there exists a limit element P (i.e., an element P with the property $\lim_{n\to\infty} \overline{P_n P} = 0$).

(3) $\mathfrak{S}$ contains a denumerable subset that is dense in $\mathfrak{S}$.

(4) $\mathfrak{S}$ has no isolated points.

The reader will have no difficulty in introducing in Section 2 the slight modifications that are requisite for making the definitions and the reasoning applicable to a metric, complete space $\mathfrak{S}$ with a dense, denumerable subset and without isolated points. By way of illustration, we show * that the distinction between exhaustible and residual sets may always be made for such a space. For suppose $M_1, M_2, \cdots, M_n, \cdots$ are all non-dense subsets of $\mathfrak{S}$. Let $\epsilon_1, \epsilon_2, \epsilon_3, \cdots$ be a sequence of positive numbers with $\lim \epsilon_n = 0$. Since M_1 is non-dense in $\mathfrak{S}$, its complement $\bar{M}_1 = \mathfrak{S} - M_1$ contains a point A_1 which is in the interior of $\bar{M}_1$ (i.e., which is not a limit point of M_1). Let C_1 be a "sphere" † with center A_1 and radius $< \epsilon_1$ containing only points of $\bar{M}_1$. C_1 contains in turn an interior point A_2 of $\bar{M}_2 = \mathfrak{S} - M_2$. Let C_2 be a sphere with center A_2 and radius $< \epsilon_2$ lying entirely in C_1 and containing only points of $\bar{M}_2$. We thus define the sequence of spheres C_n $(n = 1, 2, \cdots)$. The sequence of points $A_1, A_2, \cdots$ is evidently a regular sequence, since $A_\lambda, \lambda > n$, lies in the sphere C_n, hence its distance from A_n is less than ϵ_n, and therefore according to property (1) of $\mathfrak{S}$ the distance between A_λ and A_μ $(\lambda > n, \mu > n)$ is less than 2ϵ. Therefore, according to property (2) of $\mathfrak{S}$ there exists a point A of $\mathfrak{S}$ which is the limit of A_n.‡ A lies in every C_n, hence outside of every M_n. Therefore $\mathfrak{S}$ cannot be represented as the sum of a denumerable set of non-dense sets; that is, $\mathfrak{S}$ is not exhaustible. Since the sum of two exhaustible sets is exhaustible, it follows that a residual set, which is the complement in $\mathfrak{S}$ of an exhaustible set, is not itself exhaustible.

Further details of the extension of the ideas of the preceding section to $\mathfrak{S}$ we leave to the reader. We may state the following

THEOREM IV. *Let $\mathfrak{S}$ be any complete, metric space without isolated points and containing a dense, denumerable subset; and $f(P)$ any real function defined for the elements P of $\mathfrak{S}$.* § *Then there exists a residual set R such that if P is a point*

* Cf., for example, Hausdorff, loc. cit., chap. VIII, Section 9, especially Theorem VI and p. 328.

† A "sphere" of $\mathfrak{S}$ with center P and radius r is the set of points of $\mathfrak{S}$ whose écart from P is $\leqq r$.

‡ It is to be observed that if the radius of C_n does not approach 0, no conclusion with reference to the existence of a limit point of $\{A_n\}$ can be made. For this sequence would then not necessarily form a regular sequence; there may therefore be no limit point unless $\mathfrak{S}$ were assumed to be compact. In this case, however, function space would not be an instance for our results, since a "sphere" of function space, while possessing the properties demanded of $\mathfrak{S}$, is not compact.

§ In the terminology of E. H. Moore, f is a function on $\mathfrak{S}$ to A, where A is the set of real numbers.

of R, and N_ϵ a partial neighborhood of P, the function f is inexhaustibly, and therefore densely, approached at P via RN_ϵ. Also there exists a dense subset D of $\mathfrak{S}$ such that $f(P)$ is continuous if P ranges over D.

As particular examples of a complete metric space with a dense denumerable subset and without isolated points, we mention:

(*a*) Euclidean n-space where the écart between two points is the euclidean distance between them.

(*b*) A perfect subset of a euclidean space.

(*c*) Hilbert space, that is, the ensemble $\mathfrak{S}$ of sequences $(x_1, x_2, \cdots, x_n, \cdots)$ of real numbers with convergent $\sum_{n=1}^{\infty} x_n^2$. The écart between two "points" $(x_1, x_2, \cdots, x_n, \cdots)$ and $(y_1, y_2, \cdots, y_n, \cdots)$ is defined to be

$$\sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \cdots}.$$

The denumerable set consisting of all the sequences $(r_1, r_2, \cdots, r_n, 0, 0, 0, \cdots)$, where the r's are rational numbers, is dense in $\mathfrak{S}$. Furthermore, $\mathfrak{S}$ is evidently complete. For if P', P'', $\cdots$, $P^{(n)}$, $\cdots$ is a regular sequence of points of $\mathfrak{S}$, and $P^{(n)} = (x_1^{(n)}, x_2^{(n)}, \cdots)$, then for $\epsilon > 0$ an integer i exists such that

$$\sqrt{\sum_{n=1}^{n=\infty} (x_n^{(p)} - x_n^{(q)})^2} < \epsilon$$

for $p > i$ and $q > i$. Hence $x_n^{(p)}$ converges to x_n, say; therefore, $P^{(n)}$ to $(x_1, x_2, \cdots, x_n, \cdots)$, which, as may be seen, again belongs to Hilbert space. $\mathfrak{S}$, of course, has no isolated points.

(*d*) Function space: $\mathfrak{S}$ consists of all real continuous functions $f(x)$ defined for $0 \leqq x \leqq 1$. The écart between $f_1(x)$ and $f_2(x)$ is defined to be $\max |f_1(x) - f_2(x)|$; the postulates for écart are then satisfied. That there is a dense, denumerable subset follows from the theorem of Weierstrass that every continuous function is the limit of a uniformly convergent sequence of polynomials, which may be assumed to have rational coefficients and are thus denumerable. Moreover, since the limit of a uniformly convergent sequence of continuous functions is continuous, it follows that a regular sequence has a limit, and therefore $\mathfrak{S}$ is complete. $\mathfrak{S}$ obviously has no isolated points.

The assumption that f is single-valued may also be dropped without invalidating Theorems I and II of the last section; Theorem III, of course, implies single-valuedness by its very nature. We thus get the following results:

THEOREM I$^{(a)}$. *Let $f(x, y)$ be any real function defined for the entire XY plane and taking at every point at least one value; the number of values may change, however, from point to point and vary from 1 to c, the cardinal number of the continuum. Then the points (x, y) such that every surface point $(x, y,$*

$f(x, y)$) is densely approached by the surface $z = f(x, y)$ constitute an exhaustible set.

THEOREM II$^{(a)}$. *If $f(x, y)$ is any real, single- or many-valued function defined for the entire XY plane, there exists in the XY plane a residual set R, such that if (x, y) is any point of R, and $N_<$ a partial neighborhood of (x, y), then every point $(x, y, f(x, y))$ is inexhaustibly approached via $RN_<$.*

The proof for these more general results is essentially the same as that for Theorems I and II and is left to the reader.

Theorems I$^{(a)}$ and II$^{(a)}$ hold also for any complete, metric space with a dense, denumerable subset and without isolated points.

3. Metric properties

As in the case of the descriptive properties of Section 1, we shall confine the discussion in this section to planar sets and to one-valued functions of two variables.

Let S be any planar set; P, a point of S; C_r, a circle (interior and boundary) with P as center and r as radius; $m(C_r)$, the area of C_r, and $m_e(SC_r)$, the exterior Lebesgue measure of the portion of S in C_r. Then if

$$\lim_{r \to 0} \frac{m_e(SC_r)}{m(C_r)}$$

exists and is equal to l, we shall say that the *exterior metric density* of S at the point P is l. We have the following

THEOREM V.* *Let S be any planar set. Then the points of S at which the exterior metric density of S is $\neq 1$, i.e., the points where the exterior metric density either does not exist or does exist and is < 1, constitute a set of zero measure (Lebesgue).*

Proof. If P is a point at which the exterior metric density of S does not exist or exists and $\neq 1$, the

$$\liminf_{r \to 0} \frac{m_e(SC_r)}{m(C_r)}$$

must be < 1. Let T_k be the totality of points of S where

$$\liminf_{r \to 0} \frac{m_e(SC_r)}{m(C_r)} < k < 1.$$

The set whose measure we are to prove equal to zero is the sum of $T_{1/2}$, $T_{3/4}$, $\cdots$, $T_{(n-1)/n}$, $\cdots$; and since the sum of $\aleph_0$ sets each of zero measure is again

* For the case where only measurable sets are admitted, cf. Lebesgue, *Leçons sur l'Intégration*, 1904, pp. 124–125; Denjoy, Journal de Mathématiques, ser. 7, vol. 1 (1915), p. 132; Lusin and Sierpinski, Rendiconti del Circolo Matematico di Palermo, vol. 42 (1917), p. 167; de la Vallée Poussin, *Cours d'Analyse*, vol. 2, 1912, p. 114. For the linear case of general (not necessarily measurable) sets, cf. Blumberg, Bulletin of the American Mathematical Society, vol. 25 (1919), p. 350.

of zero measure, it is sufficient to prove that T_k is of zero measure. Let m be the exterior measure of T_k so that T_k may be enclosed, in such a way that every point of T_k is an interior point of at least one C_n, in a sequence of circles C_n, $n = 1, 2, \cdots$, with the sum of their areas $< m + \epsilon$, where ϵ is arbitrarily small. Associate with every point P of T_k a circle C_P lying in the interior of one of the circles C_n, and such that

$$\frac{m_e(C_P S)}{m(C_P)} < k$$

and, a fortiori,

$$\frac{m_e(C_P T_k)}{m(C_P)} < k,$$

where

$$m(C_P) = \text{area of } C_P;$$

this is possible because of the assumed property of the points of T_k. From the circles C_P a denumerable number may be extracted having the same totality of interior points, and from this denumerable number, a finite number C', C'', $\cdots$, $C^{(h)}$, that cover T_k except for a set of exterior measure $< \epsilon$. We now select from the $C^{(n)}$ a subset of non-overlapping circles as follows. Let D' be a $C^{(n)}$ having the maximum area attained by the $C^{(n)}$; remove from the set $\{C^{(n)}\}$ the circles which intersect D', and let D'' be one of the circles left, and having the maximum area for these circles; remove now all circles intersecting D'', and let D''' be one of the circles now remaining and having the maximum area for all circles intersected neither by D' nor by D''. In this way we obtain a set of non-overlapping circles $D^{(n)}$. Let a_1 represent the area of the portion of the plane covered by one or more $C^{(n)}$, and a_2 the portion of the plane covered by the $D^{(n)}$; then it follows from the manner in which the $D^{(n)}$ were chosen that

$$a_2 > \frac{1}{9} a_1.$$

Since every $D^{(n)}$ is a $C^{(n)}$ we have

$$\frac{m_e(D^{(n)} T_k)}{m(D^{(n)})} < k;$$

and since the $D^{(n)}$ do not overlap, we may cover the subset of T_k that lies in the $D^{(n)}$ by means of circles having a total area $< ka_2$. We may thus cover T_k by means of circles having as sum of their areas a number less than $ka_2 + (a_1 - a_2) + \epsilon$, the first term representing a sufficient amount for the portion of T_k in the $D^{(n)}$, the second term, for the portion of T_k in the $C^{(n)}$ and outside of the $D^{(n)}$, and the third term, for the portion of T_k outside of the $C^{(n)}$. Therefore,

$$m < ka_2 + (a_1 - a_2) + \epsilon = a_1 - (1 - k) a_2 + \epsilon;$$

and in virtue of the inequalities

$$m - \epsilon < a_1 < m + \epsilon, \qquad a_2 > \frac{1}{9} a_1,$$

we have

$$m < m + \epsilon - \frac{1}{9}(1 - k)(m - \epsilon) + \epsilon,$$

whence

$$m < \frac{(19 - k)}{1 - k} \epsilon,$$

and therefore $m = 0$.

Definition. $N_<$ is said to be a *non-vanishing* partial neighborhood of P, if the exterior metric density of $N_<$ at P is $\neq 0$.

We have the following lemma which corresponds to Lemma I in the case of the descriptive properties.

LEMMA II. *Let $\mathfrak{R}$ be a closed relation as in Lemma* I. *The points P for which (a) $N\mathfrak{R}P$ for every neighborhood of P and (b) a non-vanishing partial neighborhood $N_<$ exists such that $N_<\mathfrak{R}P$ (i.e., $N_<\mathfrak{R}P$ is false) constitute a set of zero measure.*

Proof. Let T be the totality of points P having the properties (*a*) and (*b*). We may then show as in the proof of Lemma I that $N_<\mathfrak{R}B$ for every point B of $NN_<$, which, if $N_<$ is a non-vanishing neighborhood of P, is also a non-vanishing neighborhood of P. Every point P of T thus has a non-vanishing neighborhood every point of which does not belong to T; therefore at no point of T is the exterior metric density of T equal to 1; T is therefore of zero measure.

Let $z = f(x, y)$ be a given surface; $P = (x, y)$, a point of the XY plane; and the relation $\mathfrak{R}_{r_1r_2}$, $r_1 < r_2$ (as in the case of Section 1), such that $I\mathfrak{R}_{r_1r_2}P$ when and only when there is a point (x, y, ζ), $r_1 \leqq \zeta \leqq r_2$, which is a limit of I'. It follows as before that $\mathfrak{R}_{r_1r_2}$ is closed. Therefore, by means of Lemma II, we may conclude that the points P for which (*a*) $N\mathfrak{R}_{r_1r_2}P$ for every neighborhood N of P, and (*b*) $\overline{N\mathfrak{R}_{r_1r_2}P}$ for some non-vanishing partial neighborhood of P constitute a set $T_{r_1r_2}$ of zero measure. Let T be the sum of all the sets $T_{r_1r_2}$, as r_1 and r_2, $r_1 < r_2$, vary over all the rational numbers; T, being the sum of $\aleph_0$ sets of measure zero, is itself of measure zero. Now a point of the surface $z = f(x, y)$ is either isolated (and then (x, y) belongs to a denumerable set of points and hence to a certain set of measure zero) or else it is a limit point of the surface, and then $N\mathfrak{R}_{r_1r_2}P$ for every neighborhood N of P, if $r_1 < z < r_2$. We therefore have the following

THEOREM VI. *Let $f(x, y)$ be any real, one-valued function defined in the entire plane. Then there exists in the XY plane a set Z, dependent on f, of measure zero, such that if (a) (x, y) is any point of the XY plane not belonging*

to Z; (b) $N_<$, any non-vanishing partial neighborhood of (x, y); and (c) S, any sphere with $(x, y, f(x, y))$ as center; then there is at least one point of the surface $z = f(x, y)$ lying in the sphere S and having as projection upon the XY plane a point in $N_<$.

That the theorem becomes false if we omit the restriction that the partial neighborhood $N_<$ shall be non-vanishing, is seen from the following example. Let A be a planar non-dense perfect set of positive measure, and B its complement. Let $f(x, y) = 0$ in A and 1 in B. Since A is non-dense and closed, B is an everywhere-dense I-region and therefore every point (x, y) of A has as one of its partial neighborhoods a set $N_<$, consisting exclusively of points of B. For such an $f(x, y)$, (x, y) and $N_<$, however, the assertion of the theorem is false, although, since A is not of zero measure, there are points of A not belonging to the alleged set Z of the theorem. Of course, in virtue of Theorem V, the points of A, with the exception of those belonging to a set of measure zero, have none but vanishing partial neighborhoods that consist entirely of points of B.

It is obvious, as, indeed, the last example shows, that the points at which a function is not densely approached need not be of zero measure. If in passing, however, we alter the meaning of "densely approached" in a rather natural way, we may conclude, according to Theorem VI, that f is densely approached everywhere except at the points of a set of measure zero. The new definition of dense approach is as follows: f is said to be densely approached at P if for every $\epsilon > 0$ and for every *non-vanishing* partial neighborhood $N_<$ of P, there is in $N_<$ a point different from P where f differs from $f(P)$ by less than ϵ.

Let now the relation $\mathfrak{R}_{r_1 r_2}$, $r_1 < r_2$, be such that $I\mathfrak{R}_{r_1 r_2} P$ when and only when there is a point (x, y, ζ), $r_1 \leqq \zeta \leqq r_2$, inexhaustibly approached via I; that is to say, for every $\epsilon > 0$ and for every circle C of the XY plane with P as center, there is an inexhaustible set of points in CI at which f differs from ζ by less than ϵ. Then it follows that $\mathfrak{R}_{r_1 r_2}$ is closed. If f is inexhaustibly approached at P and a non-vanishing partial neighborhood $N_<$ of P exists via which f is not inexhaustibly approached at P, then there exist two rational numbers r_1 and r_2, $r_1 < f(P) < r_2$, such that $N\mathfrak{R}_{r_1 r_2} P$ for every neighborhood N of P and $N_< \mathfrak{R}_{r_1 r_2} P$. Therefore according to Lemma II, P belongs to a certain set $Z_{r_1 r_2}$ of measure zero associated with the closed relation $\mathfrak{R}_{r_1 r_2}$. The sum of the $Z_{r_1 r_2}$ for all possible rational numbers $r_1 < r_2$ is also of measure zero; we thus have

THEOREM VII. *The set of points P of the XY plane, at which f is inexhaustibly approached and for which a non-vanishing partial neighborhood exists via which f is exhaustibly approached at P, constitute a set of measure zero.*

It is, of course, not true that the points where f is exhaustibly approached form a set of measure zero. Because of this fact the result for inexhaustible

approach is not as pleasing as that for the case we shall consider next. We shall say that *f is neglectably approached at the point P*, if a sphere S exists with P' as center such that the projection upon the XY plane of the points of the surface that lie in S constitutes a set of zero measure; f is said to be *neglectably approached at the point P via $N_<$* if a sphere S exists with P' as center such that the projection upon the XY plane of $N'_< S$ has in common with $N_<$ a set of measure zero. We have the following

THEOREM VIII. *Let $z = f(x, y)$ be any real, single-valued function defined in the entire XY plane. Then the points P of the XY plane that possess a non-vanishing partial neighborhood via which f is neglectably approached at P constitute a set of zero measure.*

The proof of this theorem, which is a generalization of Theorem VI, is analogous to that of the latter. A few indications will therefore suffice. We note first that the set of points where f is neglectably approached is of measure zero; the proof is similar to that of the exhaustibility of the set of points at which f is exhaustibly approached.* We then define $I\Re_{r_1r_2} P$, $P \equiv (x, y)$, as follows: there is a point (x, y, ζ), $r_1 \leqq \zeta \leqq r_2$, such that if S is any sphere with (x, y, ζ) as center, and the surface points in S are projected upon the XY plane, the portion falling in I forms a set of positive exterior measure; i.e., (x, y, ζ) is not neglectably approached via I. It will be seen that $\Re_{r_1r_2}$ is closed. The rest of the argument is left to the reader.

Theorem VIII asserts that certain sets are of positive exterior measure. We shall go a little further by considering the question of exterior metric density. Let P be any point and $M_{P\epsilon}$ the set of points Q for which

$$|f(Q) - f(P)| < \epsilon.$$

We shall say that f is *quasi-continuous*† at P if for every ϵ the set $M_{P\epsilon}$ has 1 as exterior metric density at P. We have the following theorem, which generalizes Theorem VIII:

THEOREM IX. *f is quasi-continuous except at the points of a set of measure zero.*

Proof. Let $S_{r_1r_2}$, $r_1 < r_2$, be the set of points (x, y) for which

$$r_1 < f(x, y) < r_2,$$

and $T_{r_1r_2}$ be the set of points where the exterior metric density of $S_{r_1r_2}$ is not 1; $T_{r_1r_2}$ is therefore of measure zero according to Theorem V. If r_1 and r_2 are rational, we have in all $\aleph_0$ sets $T_{r_1r_2}$, and thus the points in all the $T_{r_1r_2}$ constitute a set Z of zero measure. If P is a point at which f is not quasi-continuous, then, according to definition, there exists a positive number ϵ such that

* See Section 1.

† Cf. Denjoy, Bulletin de la Société Mathématique de France, vol. 43 (1915), p. 165.

$M_{f\epsilon}$ is not of exterior metric density 1 at P. Let r_1 and r_2 be two rational numbers such that

$$f(P) - \epsilon < r_1 < f(P) < r_2 < f(P) + \epsilon.$$

It follows that P belongs to $T_{r_1 r_2}$ and therefore to Z.

Suppose now that $N_<$ is a partial neighborhood of P such that if C_r is a circle with P as center and r as radius then

$$\liminf_{r \to 0} \frac{m(C_r N_<)}{m(C_r)} = k > 0;$$

the symbol m denotes Lebesgue measure, which $N_<$, as an I-region, possesses. We shall then say that $N_<$ is a "*properly* non-vanishing partial neighborhood" of P. Suppose further that a positive ϵ exists such that

$$\liminf_{r \to 0} \frac{m_e(C_r N_< M_{f\epsilon})}{m(C_r N_<)} = l < 1.$$

We shall then say that the exterior metric density of $M_{f\epsilon}$ is $\neq 1$ at P via $N_<$. It follows then that the surface $z = f(x, y)$ is not quasi-continuous at P. For if δ is any positive number, there exists a circle C_r with arbitrarily small radius and with P as center such that

$$m(C_r N_<) > (k - \delta)\, m(C_r),$$

and

$$m_e(C_r N_< M_{f\epsilon}) < (l + \delta)\, m(C_r N_<).$$

Therefore

$$\begin{aligned} m_e(C_r M_{f\epsilon}) &= m_e(C_r N_< M_{f\epsilon}) + m_e(C_r \{C_r - N_<\} M_{f\epsilon}) \\ &< [(l + \delta)(k - \delta) + 1 - (k - \delta)]\, m(C_r) \\ &= [1 - (1 - l - \delta)(k - \delta)]\, m(C_r) \end{aligned}$$

for sufficiently small δ; $C_r - N_<$ means here the portion of C_r not in $N_<$. Since $l < 1$, $k > 0$, and δ and r may be made as small as we please, we conclude that the relative exterior measure of $M_{f\epsilon}$ is not 1 at P; that is, that f is not quasi-continuous at P. The totality of such points P is therefore of measure zero. We thus have

THEOREM X. *If P is a point not belonging to a certain set of measure zero, then for every positive ϵ the exterior metric density of $M_{f\epsilon}$ is unity at P via any properly non-vanishing partial neighborhood of P.*

Since an angular region at P, i.e., the portion of the XY plane bounded by two half lines radiating from P, is a properly non-vanishing partial neighborhood of P, it follows as a corollary of Theorem X that, if we neglect a set of measure zero, *the exterior metric density of $M_{f\epsilon}$ is, for every $\epsilon > 0$, equal to unity at P via every angular region with P as vertex.*

Concluding remarks

As in the case of the descriptive properties, it may be seen that the metric theorems may be extended to many-valued functions. Theorem VIII, for example, when thus generalized, would read as follows: *Let $z = f(x, y)$ be any real, single- or many-valued function defined in the entire XY plane. Then the points (x, y) of the XY plane for which a surface point $(x, y, f(x, y))$ and a non-vanishing partial neighborhood $N_<$ exist such that $(x, y, f(x, y))$ is neglectably approached via $N_<$ constitute a set of zero measure.*

It is evident that the metric properties hold for functions of a single variable and, in general, for functions of n variables. Extension to $\aleph_0$-space and to function space would require a satisfactory definition of measure for such spaces;* it is not our purpose to enter upon such questions here. Likewise we shall not attempt to define, by means of postulates, a general space for which our metric results are to hold.

It may be remarked that instead of projecting the surface points of $z = f(x, y)$ upon the XY plane we may project them upon the X-axis and thus obtain other properties. For example, let us define the relationship $\Re_{r_1 r_2 r_3 r_4}$ (between I-regions and points of the X-axis) as follows: $I \Re_{r_1 r_2 r_3 r_4} \xi$ if the surface points having x-coördinates in I have a limit point in the rectangle $x = \xi$, $r_1 \leqq y \leqq r_2$, $r_3 \leqq z \leqq r_4$. $\Re_{r_1 r_2 r_3 r_4}$ is evidently closed. If the r's are taken to be rational, the number of possible relations $\Re_{r_1 r_2 r_3 r_4}$ is $\aleph_0$, and we may, by the aid of Lemma II and an argument repeatedly employed in this paper, obtain the following result: *Let $z = f(x, y)$ be any single- or many-valued function defined in the entire XY plane. Let ξ be a point of the X-axis of the following character: a surface point (ξ, η, ζ) and a partial non-vanishing (linear) neighborhood $N_<$ of ξ exist such that (ξ, η, ζ) is not a limit point of surface points having x-coördinates in $N_<$. The totality of points ξ is of measure zero.*

Similar results may be obtained for other metric properties and also in the case of the descriptive properties. In the case of a function of n variables, we may project upon an $(n - 1)$-space, an $(n - 2)$-space, etc.

* In this connection cf. Gâteaux, Bulletin de la Société Mathématique de France, vol. 47 (1919), p. 47.

University of Illinois,
Urbana, Ill.

A FUNDAMENTAL SYSTEM OF INVARIANTS OF A MODULAR GROUP OF TRANSFORMATIONS*

BY

JOHN SIDNEY TURNER

1. **Introduction.** Let G be any given group of g homogeneous linear transformations on the indeterminates $x_1, \cdots, x_n$, with integral coefficients taken modulo m. Hurwitz† raised the question of the existence of a finite fundamental system of invariants of G in the case where m is a prime p, and obtained an affirmative answer when g is prime to p. Dickson‡ subsequently obtained an affirmative answer for any g.

The general case presents great difficulty, owing to the fact that resolution into irreducible factors with respect to a composite modulus is not, in general, unique. The present investigation is confined to the case in which there are two indeterminates x, y, and m is the square of a prime p. The given group will be denoted by H, the notation G being retained when $m = p$. It is proved that the $p^2 + 1$ invariants

$$L^p,\ Q^p,\ pL^\alpha Q^\beta \qquad (\alpha, \beta = 0, 1, \cdots, p-1;\ \alpha, \beta \text{ not both zero}),$$

where

$$L = yx^p - xy^p, \qquad Q = (x^{p^2-1} - y^{p^2-1})/(x^{p-1} - y^{p-1}),$$

form a fundamental system of (independent) invariants of the group H.

2. Consider the group H of all linear homogeneous transformations modulo p^2:

$$x' \equiv ax + by, \qquad y' \equiv cx + dy, \qquad ad - bc \equiv 1 \pmod{p^2}, \tag{1}$$

where a, b, c, d are integers. To each transformation of H corresponds a unique transformation of the group G:

$$x' \equiv a_1 x + b_1 y, \qquad y' \equiv c_1 x + d_1 y, \qquad a_1 d_1 - b_1 c_1 \equiv 1 \pmod{p}, \tag{2}$$

where a_1, b_1, c_1, d_1 are integers. In fact, we have only to choose

$$a_1 \equiv a, \qquad \cdots, \qquad d_1 \equiv d \pmod{p}.$$

Conversely, to each transformation (2) corresponds one or more transformations (1). For, we can choose $a \equiv a_1, \cdots, d \equiv d_1 \pmod{p}$ so that

* Presented to the Society, April 15, 1922.

† Archiv der Mathematik und Physik (3), vol. 5 (1903), p. 25.

‡ *The Madison Colloquium*, Lect. III.

$ad - bc \equiv 1 \pmod{p^2}$. For example, if $a_1 \not\equiv 0 \pmod{p}$ we may take $a \equiv a_1$, $b \equiv b_1$, $c \equiv c_1 \pmod{p}$, and determine d by $ad - bc \equiv 1 \pmod{p^2}$; evidently $d \equiv d_1 \pmod{p}$.

Hence if we reduce all the transformations of H modulo p, we obtain all the transformations of G.

3. **Definition.** A rational and integral invariant of H is a polynomial $I(x, y)$ in x and y with integral coefficients, which remains unchanged modulo p^2 under every transformation (1). That is,

$$I(x', y') \equiv I(ax + by, cx + dy) \equiv I(x, y) \pmod{p^2} \tag{3}$$

for all integers $a, \cdots, d$ such that $ad - bc \equiv 1 \pmod{p^2}$.

Evidently any rational and integral invariant is a sum of homogeneous invariants; hence we restrict the investigation to the latter.

4. THEOREM I. *Let $I(x, y)$ be a rational and integral invariant of H, and let $I_1(x, y)$ be the polynomial obtained from $I(x, y)$ by replacing each coefficient by its positive or zero residue modulo p. Then $I_1(x, y)$ will be a rational and integral invariant of G.*

We have (3) for all transformations of H. Now

$$I(x, y) \equiv I_1(x, y) \pmod{p}$$

and

$$I(ax + by, cx + dy) \equiv I_1(ax + by, cx + dy)$$
$$\equiv I_1(a_1 x + b_1 y, c_1 x + d_1 y) \pmod{p},$$

hence

$$I_1(x', y') \equiv I_1(a_1 x + b_1 y, c_1 x + d_1 y) \equiv I_1(x, y) \pmod{p}, \tag{4}$$

and by § 2 this is true for all transformations of G.

5. Now (*Madison Colloquium*, pp. 34–38),

$$I_1(x, y) \equiv k T_1^{\alpha_1} T_2^{\alpha_2} \cdots T_i^{\alpha_i} \cdots T_r^{\alpha_r} * \pmod{p},$$

where k is an integer,

$$T_1 = L, \qquad T_2 = Q, \qquad T_i = R_i(L^{\frac{1}{2}p(p-1)}, \dagger\ Q^{\frac{1}{2}(p+1)}) \qquad (i = 3, 4, \cdots, r),$$

R_i being a polynomial in its two arguments, with integral coefficients; moreover the T_i $(i = 1, 2, \cdots, r)$ contain no multiple factors, and are relatively prime modulo p. Hence

$$I(x, y) \equiv k T_1^{\alpha_1} T_2^{\alpha_2} \cdots T_i^{\alpha_i} \cdots T_r^{\alpha_r} + pF(x, y) \pmod{p^2}, \tag{5}$$

where $F(x, y)$ denotes a polynomial in x, y with integral coefficients.

* In the discussion which follows, if any α_i is zero the corresponding T_i is to be suppressed.

† If $p = 2$, we omit the divisor 2 in the exponents.

6. Discussion of equation (5). Apply to $I(x, y)$ the transformation

$$x' \equiv x + py, \qquad y' \equiv y \pmod{p^2}, \tag{6}$$

expand by Taylor's Theorem, and denote the partial derivative of T_i with respect to x by T'_i. Then

$$\begin{aligned} I(x + py, y) \equiv I(x, y) + pyk T_1^{\alpha_1 - 1} \cdots T_i^{\alpha_i - 1} \\ \times \cdots T_r^{\alpha_r - 1} \sum_{i=1}^{r} \alpha_i T_1 \cdots T_{i-1} T'_i T_{i+1} \cdots T_r \pmod{p^2}. \end{aligned} \tag{7}$$

Since (6) is a transformation of H,

$$I(x + py, y) \equiv I(x, y) \pmod{p^2}.$$

Hence either $k \equiv 0 \pmod{p}$, in which case the right member of (5) reduces to its second term, or

$$\sum_{i=1}^{r} \alpha_i T_1 \cdots T_{i-1} T'_i T_{i+1} \cdots T_r \equiv 0 \pmod{p}. \tag{8}$$

Let $(g_i, 1)$ be a point at which $T_i(x, y)$ vanishes. Then, for $j \neq i$, $T_j(x, y)$ cannot vanish at $(g_i, 1)$; for, in that event, $T_j(x, y)$ would be a factor of $T_i(x, y)$ modulo p,* contrary to § 5. Therefore from (8) we have

$$\alpha_i T'_i(g_i, 1) \equiv 0 \pmod{p}. \tag{9}$$

Hence either $\alpha_i \equiv 0$, or $T'_i(g_i, 1) \equiv 0 \pmod{p}$. In the latter case, by a known theorem on Galois imaginaries, $T_i(x, 1)$ and $T'_i(x, 1)$ have a common factor with integral coefficients modulo p. But (§ 5) $T_i(x, 1)$ contains no multiple factor modulo p. Therefore† $T'_i(x, 1) \equiv 0 \pmod{p}$, whence

$$T'_i(x, y) \equiv 0 \pmod{p}. \tag{10}$$

Hence we have

THEOREM II. *In equation* (5), *for each* $i = 1, \cdots, r$, *either* α_i *is a multiple of* p, *or* $T'_i(x, y) \equiv 0 \pmod{p}$.

COROLLARY 1. $\alpha_1 \equiv 0 \pmod{p}$.

For $T_1 = yx^p - xy^p$; hence $T'_1 = pyx^{p-1} - y^p \equiv\equiv 0 \pmod{p}$.

COROLLARY 2. $\alpha_2 \equiv 0 \pmod{p}$.

For

$$T_2 = x^{p(p-1)} + x^{(p-1)^2} y^{p-1} + \cdots + x^{p-1} y^{(p-1)^2} + y^{p(p-1)},$$

hence $T'_2 \equiv\equiv 0 \pmod{p}$.

COROLLARY 3. *If* $\alpha_i = p\beta_i$ *for* $i > 2$,

$$T_i^{\alpha_i} \equiv S_i(L^p, Q^p) \pmod{p}, \tag{11}$$

where S_i *is a polynomial in its arguments, with integral coefficients.*

* *The Madison Colloquium*, p. 38.

† Dickson, *Lecture Notes on Double Modulus and Galois Imaginaries*, § 5.

For if we expand

$$T_i^{\alpha_i} = [R_i(L^{\frac{1}{2}p(p-1)}, Q^{\frac{1}{2}(p+1)})]^{p\beta_i},$$

we observe that in each term the exponent of L is a multiple of p and that either the exponent of Q or the coefficient of the term is a multiple of p.

7. **Discussion of** $T_i'(x, y) \equiv 0 \pmod{p}$. Write

$$T_i(x, y) = \sum_{r=0}^{n} A_r l^{n-r} q^r, \tag{12}$$

where $l = L^{\frac{1}{2}p(p-1)}$,* $q = Q^{\frac{1}{2}(p+1)}$, and the coefficients A_r are integers; then

$$T_i'(x, y) = \sum_{r=0}^{n-1} A_r(n-r) l^{n-r-1} l' q^r + \sum_{r=1}^{n} A_r r l^{n-r} q^{r-1} q' \equiv 0 \pmod{p},$$

where l', q' denote the partial derivatives of l, q with respect to x. Evidently $l' \equiv 0$, $q' \not\equiv 0 \pmod{p}$; therefore

$$\sum_{r=1}^{n} r A_r l^{n-r} q^{r-1} \equiv 0 \pmod{p}. \tag{13}$$

Each term of (13) is the product of the preceding by cq/l, c a constant; the degree in x of q/l is $\frac{1}{2}(p^2 - p)$,† hence the degrees in x of the successive terms increase by $\frac{1}{2}(p^2 - p)$. Equating coefficients of x, we find in succession

$$nA_n \equiv 0, \quad \cdots, \quad rA_r \equiv 0, \quad \cdots, \quad A_1 \equiv 0 \pmod{p}.$$

Hence in each term of $\sum_{r=1}^{n} A_r l^{n-r} q^r$, either the coefficient A_r or the exponent of q is a multiple of p, and we have

THEOREM III. *If* $T_i'(x, y) \equiv 0 \pmod{p}$, *then*

$$T_i(x, y) \equiv S_i(L^p, Q^p) \pmod{p}, \tag{14}$$

where S_i *denotes a rational and integral function of its arguments, with integral coefficients.*

COROLLARY. $[T_i(x, y)]^{\alpha_i}$ *is a polynomial in* L^p, Q^p, *with integral coefficients, modulo* p.

8. THEOREM IV. L^p *is invariant under the group* H.

Write $L(x', y') = e$, $L(x, y) = f$, where x', y' are derived from x, y by any transformation (2) of the group G; then‡ $e - f \equiv 0 \pmod{p}$. Hence $e - f \equiv 0 \pmod{p}$ for every transformation (1) of H. For if, as in § 2, we choose $a_1 \equiv a, \cdots, d_1 \equiv d$ modulo p, we have

$$L(ax + by, cx + dy) \equiv L(a_1 x + b_1 y, c_1 x + d_1 y) \equiv L(x, y) \pmod{p}.$$

* If $p = 2$, we omit the divisor 2 in the exponents.

† If $p = 2$, the degree is 2.

‡ *The Madison Colloquium*, p. 35.

Also

$$e^p - f^p = (e - f + f)^p - f^p$$
$$= (e - f)[(e - f)^{p-1} + \cdots + \tfrac{1}{2}p(p-1)(e-f)f^{p-2} + f^{p-1}],$$

and each factor on the right is identically congruent to zero modulo p; hence $e^p - f^p \equiv 0 \pmod{p^2}$; that is

$$[L(ax + by, cx + dy)]^p \equiv [L(x, y)]^p \pmod{p^2}$$

for every transformation of H.

COROLLARY 1. *In the same way, it can be proved that Q^p is invariant under the group H.*

COROLLARY 2. *$pL^\alpha Q^\beta$ is invariant under the group H.*

9. THEOREM V. *Any rational and integral invariant of the group H is a rational and integral function, with integral coefficients, of the $p^2 + 1$ invariants L^p, Q^p, $pL^\alpha Q^\beta$ ($\alpha, \beta = 0, 1, \cdots, p-1$; α, β not both zero). Conversely, any such function is an invariant of H.*

In (5), the term $kT_1^{\alpha_1} T_2^{\alpha_2} \cdots T_r^{\alpha_r}$ is an invariant of H. For if any $\alpha_i \equiv 0 \pmod{p}$, then by Theorems II and IV with their corollaries, $T_i^{\alpha_i}$ is an invariant; $T_1 = L$, $T_2 = Q$, $\alpha_1 \equiv \alpha_2 \equiv 0 \pmod{p}$. While if $\alpha_j \not\equiv 0 \pmod{p}$, $T_j' \equiv 0$, and by Theorems III and IV with their corollaries $T_j^{\alpha_j}$ is an invariant.

Hence the second term $pF(x, y)$ of (5) is an invariant of H, and it follows from § 2 that $F(x, y)$ is an invariant of G. Therefore $pF(x, y)$ is the product of p by a polynomial in L and Q. It follows that if $I(x, y)$ is any rational and integral invariant of H,

$$I(x, y) \equiv S(L^p, Q^p, pL^\alpha Q^\beta) \pmod{p^2}, \tag{15}$$

where $pL^\alpha Q^\beta$ denotes the set pL, pQ, pL^2, pLQ, $\cdots$, $pL^{p-1}Q^{p-1}$, and S denotes a rational and integral function of its arguments, with integral coefficients.

Conversely, any rational and integral function of L^p, Q^p, $pL^\alpha Q^\beta$, with integral coefficients, is a sum of invariants, and is therefore itself an invariant. Hence these $p^2 + 1$ invariants form a fundamental system.

10. THEOREM VI. *The invariants of the fundamental system are independent.*

In view of the coefficients p, neither L^p nor Q^p can be expressed as a polynomial in the remaining invariants, with integral coefficients. Assume that $pL^{\alpha_1} Q^{\beta_1}$, $\alpha_1 \leqq p-1$, $\beta_1 \leqq p-1$, can be so expressed. Then

$$pL^{\alpha_1} Q^{\beta_1} \equiv P(L^p, Q^p, pL^\alpha Q^\beta) \pmod{p^2} \tag{16}$$

identically in x, y. We may suppose that P contains no group of terms which vanishes identically modulo p^2. Let $mL^{\alpha_2} Q^{\beta_2}$ be any term of P; then $pL^{\alpha_1} Q^{\beta_1}$ and $mL^{\alpha_2} Q^{\beta_2}$ must be of the same total degree in x, y, and also of the same

degree in x alone. Therefore

$$\alpha_1(p+1)+\beta_1 p(p-1)=\alpha_2(p+1)+\beta_2 p(p-1),$$
$$\alpha_1 p+\beta_1 p(p-1)=\alpha_2 p+\beta_2 p(p-1),$$

whence $\alpha_1=\alpha_2$, $\beta_1=\beta_2$. Hence P consists of the single term $pL^{\alpha_1}Q^{\beta_1}$. Evidently $pL^{\alpha_1}Q^{\beta_1}$ is not a product of fundamental invariants, hence the theorem is proved.

11. If we consider the total group

$$x'\equiv ax+by, \qquad y'\equiv cx+dy \pmod{p^2},$$
$$ad-bc\not\equiv 0 \pmod{p}, \tag{17}$$

we find, exactly as in Theorem IV, that Q^p is an absolute invariant, and that L^p, $pL^\alpha Q^\beta$ are relative invariants of indices p, α, respectively.

IOWA STATE COLLEGE,
AMES, IOWA.

THE GAUSSIAN LAW OF ERROR FOR ANY NUMBER OF VARIABLES*

BY

J. L. COOLIDGE

The exponential law for the distribution of accidental errors of observation, discovered by Gauss, has been a mathematical classic for over a century. Many have been the attempts to prove it, all based, necessarily, on more or less arbitrary assumptions. Perhaps the most searching examination of it was given by Poincaré in his *Calcul des probabilités*; his final opinion seems to be contained in the following phrase:†

"J'ai plaidé de mon mieux jusqu'ici en faveur de la loi de Gauss dont nous allons maintenant tirer les conséquences. Peut-être pourtant la cause n'était-elle pas parfaitement bonne.

"Elle ne s'obtient pas par des déductions rigoureuses, plus d'une démonstration qu'on a voulu en donner est grossière, entre autres celle qui s'appuie sur l'affirmation que la probabilité des écarts est proportionnelle aux écarts. Tout le monde y croit cependant, me disait un jour M. Lippmann, car les expérimenteurs s'imaginent que c'est un théorème de mathématiques, et les mathématiciens, que c'est un fait expérimental."

The law has been extended to include the distribution of errors depending upon two variables, and in this form it has a certain importance in the theory of ballistics, and in that of statistical correlation; even the case of three variables has been slightly treated. The general case of n variables has never been taken up except in two recent articles by von Mises.‡ The treatment here is based on a very general form of analysis showing how an arbitrary distribution function will lead asymptotically to an exponential form. The analysis is very careful, the point of view extremely abstract, with little relation to practical applications. Moreover, the author gives no indication how the constants should be calculated in any particular case. It is the object of the present paper to deduce the Gaussian law for n variables by a method based upon the classical one for a single variable, but with somewhat broader and more explicit assumptions. In the second part we shall make

* Presented to the Society, December 27, 1922.

† Poincaré, *Calcul des Probabilités*, Paris, 1896, pp. 196 and 149.

‡ *Fundamentalsätze der Wahrscheinlichkeitsrechnung*, Mathematische Zeitschrift, vol. 4 (1919), and *Grundlagen der Wahrscheinlichkeitsrechnung*, ibid., vol. 5 (1920). See also Dodd, *Functions of measurements*, Skandinavisk Aktuarietidskrift, 1922.

the additional assumptions necessary to determine the coefficients in any particular case, and show how these latter may then be calculated in a simple manner.

1. The deduction of the law

Suppose that we are concerned with measuring groups of m quantities. We shall, for simplicity, assume that all groups are equally trustworthy, although the extension to the case of differently weighted groups is not difficult. We shall make certain assumptions about the distribution of errors, meaning, thereby, *accidental errors*, for we assume that constant errors have been removed.

Assumption 1. *The a priori probability that a group of quantities to be measured should take values in the infinitesimal region*

$$X \pm \tfrac{1}{2}dX, \qquad Y \pm \tfrac{1}{2}dY, \qquad Z \pm \tfrac{1}{2}dZ, \qquad \cdots,$$

where the points $X, Y, Z, \cdots$ lie in a certain continuous m-dimensional manifold S, will differ by an infinitesimal of higher order from the expression

$$f(X, Y, Z, \cdots)\,dX\,dY\,dZ\cdots,$$

where the function f is continuous with continuous first derivatives in S.

Assumption 2. *The probability that a group of quantities whose true values are $X, Y, Z, \cdots$ in S should be observed to have values, after the removal of constant errors, which lie in the infinitesimal region*

$$x \pm \tfrac{1}{2}dx, \qquad y \pm \tfrac{1}{2}dy, \qquad z \pm \tfrac{1}{2}dz, \qquad \cdots,$$

where $(x, y, z, \cdots)$ is a point of S, will differ by an infinitesimal of higher order from

$$\Phi(X, Y, Z, \cdots, x, y, z, \cdots)\,dxdydz\cdots,$$

where the function Φ is continuous with continuous first and second partial derivatives, and has a value independent of the choice of origin.

The last part of the assumption is plausible in practice, because if we are, for instance, measuring a length on a scale, the accidental errors will arise from various physical causes independent of the position of the 0. Moreover, it has a momentous consequence, for

$$\Phi(X, Y, Z, \cdots, x, y, z, \cdots) = \Phi|0, 0, 0, \cdots, x - X, y - Y, z - Z, \cdots|.$$

Writing in the explicit values of the errors we have

$$\xi = x - X, \qquad \eta = y - Y, \qquad \zeta = z - Z,$$
$$\Phi \equiv \Phi(\xi, \eta, \zeta, \cdots).$$

that is to say, the probability for a system of errors is a function of those errors, and not of the true values and observed values considered as independent variables, a point which has been a stumbling block to some writers.

ASSUMPTION 3. *The mean value for the error on an individual variable is* 0.

This again is plausible, for a contrary assumption would show a tendency to favor positive or negative errors, and such a tendency we should naturally class with the constant errors, not with the accidental ones. As a further matter of notation let us write the averages

$$\bar{x} = \frac{x_1 + x_2 + \cdots + x_n}{n}, \qquad \bar{y} = \frac{y_1 + y_2 + \cdots + y_n}{n}, \tag{1}$$
$$\bar{z} = \frac{z_1 + z_2 + \cdots + z_n}{n}, \qquad \cdots.$$

The actual errors are

$$\xi_i = x_i - X, \qquad \eta_i = y_i - Y, \qquad \zeta_i = z_i - Z, \qquad \cdots. \tag{2}$$

The residual errors are

$$\begin{aligned} \delta_i &= x_i - \bar{x}, \qquad \epsilon_i = y_i - \bar{y}, \qquad \cdots, \\ x_i - \bar{x} &= x_i - X - (\bar{x} - X), \\ \delta_i &= \xi_i - \frac{\sum_j \xi_j}{n}. \end{aligned} \tag{3}$$

The individual groups of observations are independent of one another; hence, by Assumption 3,

$$\begin{aligned} &\text{Mean value } \xi_i \xi_j = 0, \\ &\text{Mean value } \delta_i^2 = \frac{n-1}{n} \text{Mean value } \xi_i^2, \\ &\text{Mean value } \xi_i^2 = \text{Mean value } \frac{\sum_j \delta_j^2}{n-1}, \\ &\text{Mean value } \xi_i \eta_i = \text{Mean value } \frac{\sum_j \delta_j \epsilon_j}{n-1}. \end{aligned} \tag{4}$$

ASSUMPTION 4. *If the infinitesimal increments* dx, dy, dz, $\cdots$ *be sufficiently small, the probability that the true values lie in the region*

$$\bar{x} \pm \tfrac{1}{2}dx, \qquad \bar{y} \pm \tfrac{1}{2}dy, \qquad \bar{z} \pm \tfrac{1}{2}dz, \quad \cdots$$

is greater than that they lie in any other region of like structure about any other point.

We have now made a sufficient number of assumptions to enable us to deduce the analytic form for our functions. We do this, following the original method of Gauss, by calculating the probability that a given set of observations should have resulted from observing a group of quantities of assumed true value. The probability that the measurements $x_1, y_1, z_1, \cdots, x_2, y_2, z_2, \cdots, x_n, y_n, z_n, \cdots$ were made on quantities whose true values are $X, Y, Z,$

$\cdots$ is by Bayes' theorem

$$\frac{f(X, Y, Z, \cdots)\Phi_1\Phi_2\cdots\Phi_n\, dXdYdZ\cdots}{\int\cdots\int f(X, Y, Z, \cdots)\Phi_1\Phi_2\cdots\Phi_n\, dXdYdZ\cdots},$$

$$\Phi_i = \Phi(\xi_i, \eta_i, \zeta_i, \cdots).$$

The integration in the denominator is supposed to be extended throughout the whole region S. This expression will be a maximum with the logarithm of its numerator. Equating to 0 the partial derivatives to X, Y, Z, $\cdots$, we get

$$-\frac{\partial \log f}{\partial X} + \frac{\partial \log \Phi_1}{\partial \xi_1} + \frac{\partial \log \Phi_2}{\partial \xi_2} + \cdots + \frac{\partial \log \Phi_n}{\partial \xi_n} = 0,$$

$$-\frac{\partial \log f}{\partial Y} + \frac{\partial \log \Phi_1}{\partial \eta_1} + \frac{\partial \log \Phi_2}{\partial \eta_2} + \cdots + \frac{\partial \log \Phi_n}{\partial \eta_n} = 0.$$

One set of solutions will arise in case all of the observations have been correct, i.e.,

$$-\frac{\partial \log f}{\partial X} + n\frac{\partial \log \Phi(\xi, \eta, \cdots)}{\partial \xi} = 0, \qquad \xi = \eta = \cdots = 0,$$

$$-\frac{\partial \log f}{\partial Y} + n\frac{\partial \log \Phi(\xi, \eta, \cdots)}{\partial \eta} = 0, \qquad \xi = \eta = \cdots = 0.$$

Now, by definition, f is independent of n, hence

$$\frac{\partial \log f}{\partial X} = \frac{\partial \log f}{\partial Y} = \cdots = 0, \qquad f = \text{const.}$$

Let us underline the fact that we are considering probabilities for observations which do not go outside of the region S. We could not have f a constant throughout all space without a contradiction, and it is also evident that f must be rigorously 0 throughout most of space. The partial differential equations now take the simpler form

$$\begin{aligned}
&\frac{\partial \log \Phi_1}{\partial \xi_1} + \frac{\partial \log \Phi_2}{\partial \xi_2} + \cdots + \frac{\partial \log \Phi_n}{\partial \xi_n} = 0, \\
&\frac{\partial \log \Phi_1}{\partial \eta_1} + \frac{\partial \log \Phi_2}{\partial \eta_2} + \cdots + \frac{\partial \log \Phi_n}{\partial \eta_n} = 0, \\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots
\end{aligned} \tag{5}$$

These equations hold whenever

$$\begin{aligned}
&X = \bar{x}, \qquad Y = \bar{y}, \qquad Z = \bar{z}, \qquad \cdots, \\
&\xi_1 + \xi_2 + \cdots + \xi_n = 0, \\
&\eta_1 + \eta_2 + \cdots + \eta_n = 0, \\
&\cdots\cdots\cdots\cdots\cdots\cdots
\end{aligned}$$

By Assumption 2 we are free to treat our assumed groups as if they were absolutely independent quantities, provided we do not go outside of S. Let us, then, assume that the observed groups $x_1, x_2, \cdots, x_n; y_1, y_2, \cdots, y_n; z_1, z_2, \cdots, z_n$ take such infinitesimal increments that the averages $\bar{x}, \bar{y}, \bar{z}, \cdots$ in (1) are not altered:

$$\frac{\partial}{\partial \xi_1}\left[\frac{\partial \log \Phi_1}{\partial \xi_1}\right]\delta\xi_1 + \frac{\partial}{\partial \xi_2}\left[\frac{\partial \log \Phi_2}{\partial \xi_2}\right]\delta\xi_2 + \cdots + \frac{\partial}{\partial \xi_n}\left[\frac{\partial \log \Phi_n}{\partial \xi_n}\right]\delta\xi_n = 0,$$

$$\frac{\partial}{\partial \xi_1}\left[\frac{\partial \log \Phi_1}{\partial \eta_1}\right]\delta\xi_1 + \frac{\partial}{\partial \xi_2}\left[\frac{\partial \log \Phi_2}{\partial \eta_2}\right]\delta\xi_2 + \cdots + \frac{\partial}{\partial \xi_n}\left[\frac{\partial \log \Phi_n}{\partial \eta_n}\right]\delta\xi_n = 0,$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots,$$

$$\frac{\partial}{\partial \eta_1}\left[\frac{\partial \log \Phi_1}{\partial \eta_1}\right]\delta\eta_1 + \frac{\partial}{\partial \eta_2}\left[\frac{\partial \log \Phi_2}{\partial \eta_2}\right]\delta\eta_2 + \cdots + \frac{\partial}{\partial \eta_n}\left[\frac{\partial \log \Phi_n}{\partial \eta_n}\right]\delta\eta_n = 0.$$

We have also

$$\delta\xi_1 + \delta\xi_2 + \cdots + \delta\xi_n = 0,$$
$$\delta\eta_1 + \delta\eta_2 + \cdots + \delta\eta_n = 0,$$
$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots.$$

Each of the first set of equations in the variables $d\xi_1, d\xi_2, \cdots, d\xi_n$ must hold whenever the last equation in these variables holds, hence, integrating once, and dropping subscripts,

$$\frac{\partial \log \Phi}{\partial \xi} = \alpha\xi + \beta\eta + \cdots + \omega.$$

Giving ξ the successive values $\xi_1, \xi_2, \cdots, \xi_n$, and doing the same for the other variables, and then summing, we have

$$\omega = 0,$$

$$\frac{\partial \log \Phi}{\partial \xi} = \alpha\xi + \beta\eta + \cdots,$$

(6)
$$\frac{\partial \log \Phi}{\partial \eta} = \alpha'\xi + \beta'\eta + \cdots,$$

$$\Phi \equiv Re^{-\psi^2(\xi, \eta, \zeta, \cdots)}.$$

Here the expression ψ^2 is a homogeneous quadratic form in the variables.

Our assumptions are sufficient to enable us to make a very definite statement about the function ψ^2, namely, that its discriminant is not zero. For if the discriminant were zero, the partial derivatives would be linearly dependent, and vanish for an infinite number of sets of values for the variables, and this is directly in conflict with our fourth assumption that the only maximum

arose from taking all of these variables equal to zero. Furthermore, since this is a maximum, we know that the form is definite, i.e.,

The homogeneous quadratic form ψ^2 is positive and definite, with a non-vanishing discriminant.

2. Determination of the constants*

It should be emphasized that everything which we have done so far is under the assumption that we are dealing with observations in the region S. We have found the probability that an observation in the region S should lie in a certain infinitesimal sub-region. In practice this is of no interest whatever until we have some idea of what the region S may be. It certainly could not be the whole of space, as the assumption that f is everywhere constant leads to a contradiction. On further consideration we notice two things. First of all, it seems quite plausible that f might be constant throughout a certain region, and equal to 0 almost everywhere else. Second, the expression (6) is excessively small, except in a very strictly confined space, and this rapid diminution of (6) would produce a result close to that of the vanishing of f. In other words, the error in calculating the constants will be very small if we assume that the formula (6) is universally valid. On the strength of this we make

Assumption 5. *For the purpose of calculating constants, formula* (6) *may be assumed true throughout all space.*

We note, secondly, that the only method for calculating our constants is to assume that certain observed values may be identified with their mean values as calculated by formula. The right sides of the last two equations (4) are proportional to the mean values of the averages of certain observed quantities, and we know by Tchebycheff's theorem† that it is highly likely that the value of an average shall be close to its mean value. This leads to

Assumption 6. *When the number of groups is large, the mean values of $\xi^2, \eta^2, \xi, \eta, \cdots$ may be equated to the observed values*

$$\frac{\sum_j \delta_j^2}{n-1}, \quad \frac{\sum_j \epsilon_j^2}{n-1}, \quad \frac{\sum_j \delta_j \epsilon_j}{n-1}, \quad \cdots .$$

These quantities give, of course, the probable errors of individual measure-

* The mathematical manipulation that follows depends on obvious applications of the theory of determinants. The methods and final formula are very close to Greiner, Zeitschrift für Mathematik und Physik, vol. 57, pp. 226 ff., and Pearson, Philosophical Transactions of the Royal Society, vol. 187, pp. 299 ff. Pearson assigns the credit to Edgeworth, Philosophical Magazine, ser. 5, vol. 34, p. 201. I must confess to finding Edgeworth so obscure that I do not know whether his result is like mine or not. Moreover, none of these writers seem to me to set forth the underlying assumptions with desirable clearness.

† Tchebycheff, *Oeuvres*, Petrograd, 1899, vol. 1, p. 687.

ments and their correlation coefficients two by two. It is interesting that these should be the only independent constants.

In order to clarify the manipulation, we shall at this point take the perilous step of changing our notation. Logically, the notation we are now going to adopt might well have been used from the start, but the resulting summation formulas, by their very compactness, are rather obscure, and it is easier to see what is really going on, by using the more diffuse symbolism with many continuation signs which we have employed so far. For the actual errors committed, we shall write

$$\xi = x_1, \qquad \eta = x_2, \qquad \zeta = x_3, \qquad \cdots, \tag{7}$$

there being m in all. The n sets of m residuals shall be written

$$\delta_{11}, \delta_{12}, \cdots, \delta_{1m}; \quad \delta_{21}, \delta_{22}, \cdots, \delta_{2m}; \quad \cdots; \quad \delta_{n1}, \delta_{n2}, \cdots, \delta_{nm}.$$

Our fundamental formula (6) may now be written

$$\Phi \equiv Re^{-\Sigma a_{ij}x_i x_j}, \qquad a_{ij} = a_{ji}. \tag{8}$$

The assumptions 5 and 6 may be expressed by the equation

$$p_{ij} \equiv \frac{\sum_k \delta_{ki}\,\delta_{kj}}{n-1} = R\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} x_i\,x_j\,e^{-\Sigma a_{ij}x_i x_j}\,dx_1\,dx_2\cdots dx_m. \tag{9}$$

Since the discriminant of our quadratic form is not zero, we may find a linear transformation

$$x_i = \sum_k c_{ik}\,x'_k, \qquad |c_{ij}| \neq 0, \tag{10}$$

such that

$$\sum_{i,j} a_{ij}\,x_i\,x_j \equiv \sum_{i,j,k,l} c_{ik}\,c_{jl}\,a_{ij}\,x'_k\,x'_l = \sum_r b_r\,x'^2_r,$$

$$\sum_{i,j} c_{ik}\,c_{jl}\,a_{ij} = 0, \qquad k \neq l, \tag{11}$$

$$\sum_{i,j} c_{ir}\,c_{jr}\,a_{ij} = b_r.$$

Since the discriminant is an invariant of weight 2,

$$b_1\,b_2\cdots b_m = |c_{ij}|^2\cdot|a_{ij}|.$$

The inverse of the substitution, contragredient to (10), is

$$w'_k = \sum_i c_{ik}\,w_i,$$

$$w'^2_k = \sum_{i,j} c_{ik}\,c_{jk}\,w_i\,w_j.$$

In the projective space of $n-1$ dimensions where a point has the homogeneous coördinates $x_1, x_2, \cdots, x_n$ the hyperquadric

$$\sum_{i,j} a_{ij}\,x_i\,x_j = 0$$

has the tangential equation

$$\sum_{i,j} A_{ij}\, w_i\, w_j = 0, \qquad A_{ij} = \frac{\partial |a_{ij}|}{\partial a_{ij}}.$$

In terms of the new variables we have

$$\sum_r b_r\, x_r'^2 = 0, \qquad \sum_r \frac{w_r'^2}{b_r} = 0,$$

$$|b_1\, b_2 \cdots b_m| \sum_r \frac{w_r'^2}{b_r} = |c_{ij}|^2 \sum_{i,j} A_{ij}\, w_i\, w_j, \tag{12}$$

$$|a_{ij}| \sum_r \frac{c_{ir}\, c_{jr}}{b_r} = A_{ij}.$$

We may express (9) in terms of the new variables. The only point to remember is that the jacobian of the transformation is $|c_{ij}|$;

$$p_{ij} = R|c_{ij}| \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \sum_{k,l} c_{ik}\, c_{jl}\, x_k'\, x_l'\, e^{-\sum b_r x_r'^2}\, dx_1'\, dx_2' \cdots dx_m'.$$

This simplifies greatly because

$$\int_{-\infty}^{\infty} x_k'\, e^{-b_k' x_k'^2}\, dx_k' = 0.$$

Hence

$$p_{ij} = R|c_{ij}| \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \sum_k c_{ik}\, c_{jk}\, x_k'^2\, e^{-\sum b_r x_r'^2}\, dx_1'\, dx_2' \cdots dx_m'.$$

We have the well known integrals

$$\int_{-\infty}^{\infty} e^{-b_k x_k'^2}\, dx_k' = \frac{\sqrt{\pi}}{\sqrt{b_k}}; \qquad \int_{-\infty}^{\infty} x_k'^2\, e^{-b_k x_k'^2}\, dx_k' = \frac{\sqrt{\pi}}{2b_k^{3/2}}.$$

Hence

$$p_{ij} = R|c_{ij}|\, \frac{\pi^{m/2}}{2\sqrt{b_1\, b_2 \cdots b_m}} \sum_k \frac{c_{ik}\, c_{jk}}{b_k},$$

or, by (12),

$$p_{ij} = \frac{R}{2} \frac{|c_{ij}|}{|a_{ij}|} \frac{\pi^{m/2}}{\sqrt{b_1\, b_2 \cdots b_m}} A_{ij} = \frac{R}{2} \frac{\pi^{m/2}\, A_{ij}}{|a_{ij}|^{3/2}}.$$

Furthermore, since the probability of some group of errors is 1,

$$1 = R \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} e^{-\sum a_{ij} x_i x_j}\, dx_1 \cdots dx_n,$$

$$1 = \frac{R\pi^{m/2}|c_{ij}|}{\sqrt{b_1\, b_2 \cdots b_m}} = \frac{R\pi^{m/2}}{\sqrt{|a_{ij}|}}.$$

Dividing out R, we have

$$p_{ij} = \frac{A_{ij}}{2|a_{ij}|} .$$

In these equations the quantities p_{ij} are known; we wish to find the quantities a_{ij}. We first introduce one more symbol:

$$P_{ij} = \frac{\partial |p_{ij}|}{\partial p_{ij}} .$$

Since the process of interchanging each element of a non-vanishing determinant with its cofactor is an involutory one, except for multiplication by a power of the determinant, we must have

$$a_{ij} = MP_{ij} .$$

We calculate M by a little jugglery:

$$|a_{ij}| = M^m |p_{ij}|^{m-1} ,$$

$$|p_{ij}| = \frac{|A_{ij}|}{2^m |a_{ij}|^m} = \frac{1}{2^m |a_{ij}|} ,$$

$$M = \frac{1}{2|p_{ij}|} .$$

Let us exhibit our results in the form of a table:
Assumed errors: $x_1, x_2, \cdots, x_n$.
Gaussian law of error for n variables: $\Phi = Re^{-\Sigma a_{ij} x_i x_j}$.
Observed residuals: $\delta_{11}, \delta_{12}, \cdots, \delta_{1m}; \delta_{21}, \delta_{22}, \cdots, \delta_{2m}; \cdots; \delta_{n1}, \delta_{n2}, \cdots, \delta_{nm}$.

$$p_{ij} = \sum_k \frac{\delta_{ki} \delta_{kj}}{n-1} ; \qquad P_{ij} = \frac{\partial |p_{ij}|}{\partial p_{ij}} ;$$

$$a_{ij} = \frac{P_{ij}}{2|p_{ij}|} ; \qquad R = \frac{1}{\sqrt{(2\pi)^m |p_{ij}|}} .$$

HARVARD UNIVERSITY,
CAMBRIDGE, MASS.

CERTAIN THEOREMS RELATING TO PLANE CONNECTED POINT SETS*

BY

ANNA M. MULLIKIN

I. Introduction

A point set M is said† to be connected if it cannot be expressed as the sum of two mutually exclusive point sets neither of which contains a limit point of the other. Sierpinski‡ has shown that a closed, bounded, connected set of points in space of n dimensions cannot be separated into a countable infinity of closed point sets such that no two of them have a point in common. It will be shown in the present paper that for the case where $n = 2$, this theorem does not remain true if the stipulation that M is closed be removed. It will however be shown that a plane point set, regardless of whether it be closed or bounded, which separates its plane cannot be expressed as the sum of a countable infinity of closed, mutually exclusive point sets, no one of which separates the plane. Of the other results established, the principal one is that if M_1 and M_2 are two closed, connected, bounded point sets, neither of which disconnects a plane S, a necessary and sufficient condition that their sum, M, shall disconnect S is that $\overline{M}$, the set of points common to M_1 and M_2, be not connected.

I wish to thank Professor Robert L. Moore, who suggested the theorems of this paper. Without his help and encouragement it could not have been written.

II

The following is an example of a countable collection of mutually exclusive, closed, and bounded point sets with connected sum. Consider a countable infinity of arcs each of which is made up of four straight-line intervals (Fig. 1), the nth arc being drawn from the point $(m/2^{n-1}, 0)$ to $(m/2^{n-1}, m/2^{n-1})$, thence to $(-m/2^{n-1}, m/2^{n-1})$, thence to $(-m/2^{n-1}, -m/2^{n-1})$ and thence to $(m, -m/2^{n-1})$. Let n go from one to infinity, and let M be the point

* Various parts of this paper were presented to the Society on October 25, 1919, December 28, 1920, and February 26, 1921.

† See N. J. Lennes, *Curves in non-metrical analysis situs with an application in the calculus of variations*, American Journal of Mathematics, vol. 33 (1911), and Bulletin of the American Mathematical Society, vol. 12 (1906).

‡ W. Sierpinski, *Un théorème sur les continus*, Tôhoku Mathematical Journal, vol. 13.

set composed of the sum of all the arcs so obtained. It will be seen that each of these arcs contains a limit point of every subset of M which consists of an infinite number of the remaining arcs. Hence the set M is connected. It is obviously bounded.

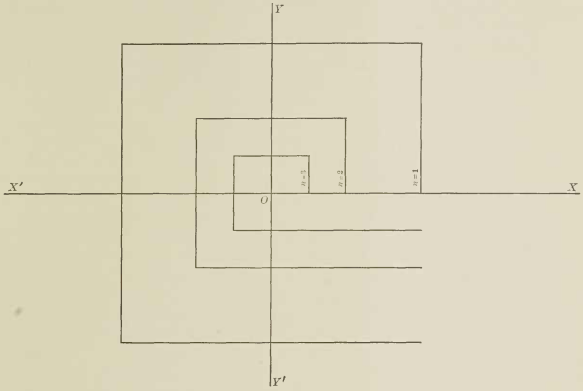

FIG. 1.

THEOREM 1.* *If, in a plane S, K and M are two closed, mutually exclusive point sets and H is a closed, bounded, connected point set having at least one point in common with each of the sets K and M, then there exists a point set $\overline{H}$, a subset of H, such that $\overline{H}$ is connected and contains no point of either K or M, but such that K and M each contain a limit point of $\overline{H}$.*

In our proof of Theorem 1, we shall make use of the following two well known theorems, A and B.

THEOREM A.† *If K and M are two closed point sets having no point in common, and H is a continuous, bounded point set having at least one point in*

* Rosenthal gives a proof for the case in which each of the sets K and M reduces to a single point. See A. Rosenthal, *Teilung der Ebene durch irreduzible Kontinua*, Sitzungsberichte der mathematisch-physikalischen Klasse der Bayerischen Akademie der Wissenschaften zu München, 1919, p. 104.

† Janiszewski gives a proof for the case in which each of the sets K and M reduces to a single point. His proof can readily be extended to the more general case. Cf. S. Janiszewski, *Sur les continus irréductibles entre deux points*, Journal de L'Ecole Polytechnique (2), vol. 16 (1912), p. 109, Theorem 1.

common with each of the sets K and M, then H contains a subset which is irreducibly continuous from K to M.

THEOREM B.* *If $H_1, H_2, H_3, \cdots$ is a countable collection of connected point sets, and P is a point such that every circle containing P contains a point from all except a finite number of these sets, then the limit set† of the sequence $H_1, H_2, H_3, \cdots$ is connected.*

Proof of Theorem 1. By Theorem A, H contains a subset which is irreducibly continuous from K to M. Call this set H_{KM}. Let $\overline{K}$ and $\overline{M}$ denote the points of H_{KM} belonging to K and M respectively. Let H' denote $H_{KM} - \overline{K} - \overline{M}$. We can show that H' is the required $\overline{H}$. Evidently it only remains to be proved that H' is connected. Consider any point P of H'. We can show that the largest connected subset of H' in which P lies has a limit point in either $\overline{K}$ or $\overline{M}$. For suppose it has not. It will then be closed and may be enclosed in a finite number of circles no one of which contains or encloses a point of either $\overline{K}$ or $\overline{M}$. The interiors of these circles form a domain D_P. Now select some point K_1 of $\overline{K}$. Since P and K_1 lie together in the connected set H_{KM} they can be joined by a broken line composed of a finite number of intervals of length less than half an inch, such that the vertices of this broken line belong to H_{KM}. Let L_1 be that vertex on this broken line which immediately precedes the first vertex on it, in the order from P to K_1, that lies without D_P. Then join P and K_1 by a broken line of intervals of length less than a quarter of an inch such that the vertices belong to H_{KM}. Let L_2 be the point on this line corresponding to L_1. Continue this process indefinitely. By Theorem B the limit set will be connected. It will contain P and a point on the boundary of D_P,‡ namely the limit point of $L_1, L_2, L_3, \cdots$. It contains only points of H_{KM}, but since it lies wholly within D_P plus its boundary, it contains no point of $\overline{K}$ or $\overline{M}$ and hence is a subset of H'. This is contrary to the hypothesis that the largest connected subset of H' in which P lies is within the domain D_P.

Denote by H_K the set of those points lying in a connected subset of H' of which $\overline{K}$ contains a limit point, by H_M the set of those which lie in a connected subset of H' of which $\overline{M}$ contains a limit point. Let $H_K + \overline{K}$ be denoted by S_K, $H_M + \overline{M}$ by S_M. Since $S_K + S_M = H_{KM}$, and since H_{KM} is connected, S_K and S_M must have a point in common or else one of these sets must contain a limit point of the other. Suppose first that they have a point in common. This point must belong to H', and it is evident that since H_{KM} is irreducibly

* See S. Janiszewski, loc. cit., p. 98, Theorem 1.

† By the limit set of a sequence of sets $H_1, H_2, H_3, \cdots$ we mean the set of all points $[P]$ such that P is a limit point of a set of points $X_1, X_2, X_3, \cdots$ such that for every k, X_k belongs to H_k.

‡ Janiszewski gives a parallel argument to prove that if the continuous set C contains a point A which is an interior point of the closed set K, then there exists a continuous set containing the point A and contained in K and C. See S. Janiszewski, loc. cit., p. 100, Theorem IV.

continuous, H' must in this case be connected. Suppose secondly that one of the sets contains a limit point of the other, for instance that S_K contains a limit point, P_K, of S_M. And suppose that H' is not connected. Since S_K contains a limit point of S_M it is evident that H_M must exist. Then if $H_M = H'$, H_M cannot be connected. But suppose that the set H_K actually exists; we can show that in this case, too, H_M is not connected. For suppose it were. $P_K + H_M$ together with the largest connected subset of H' in which P_K lies and all points in H_K which are limit points of H_M would then be a connected subset of H' and consequently a proper subset of H'. This together with its limit points in $\bar{K}$ and $\bar{M}$ would be both a continuous set between K and M and a proper subset of H_{KM}. This is contrary to the hypothesis that H_{KM} is irreducibly continuous between K and M. We have therefore shown that if H' is not connected, H_M is not connected.

Suppose this to be the case, and let $H_M = H_{M_1} + H_{M_2}$ where H_{M_1} and H_{M_2} are two mutually exclusive sets neither of which contains a limit point of the other. Suppose P_K is a limit point of H_{M_1}. Then enclose every point of H_{M_2} in a circle which encloses no point of the set $H_{M_1} + P_K + \bar{M}$. The interiors of these circles form a domain D. Now since P_K is a limit point of H_{M_1} and every point of H_{M_2} is connected with some point of $\bar{M}$ in a subset of $H_{M_1} + \bar{M}$, P_K can be joined by an infinite number of broken lines, as before, to points of $\bar{M}$ such that the vertices of these broken lines belong to $H_{M_1} + \bar{M}$, and therefore lie without D or on its boundary. The limit set will then be connected and will contain no point of D. This together with the largest connected subset of H' in which P_K lies and the limit points of this set in $\bar{K}$ will be a continuous set from K to M, a proper subset of H_{KM} since it contains no point of H_{M_2}. This is contrary to the hypothesis that H_{KM} is irreducibly continuous from K to M. We have therefore proved that H' is connected and is the required H.

THEOREM 2. *If, in a plane S, H is a closed, bounded point set containing two mutually exclusive, closed point sets K and M, but containing no closed, connected subset containing a point of K and a point of M, then it is the sum of two mutually exclusive, closed sets, of which one contains K and the other contains M.*

Proof. There exists a positive number ϵ such that no point of K can be joined to a point of M by a broken line made up of intervals of length less than ϵ such that the end points of these intervals are points of H. For otherwise there would be a closed, connected "limit set" as in Theorem 1. This limit set would belong to H, since H is closed, and it would contain a point of K and a point of M, since K and M are both closed. This is contrary to the hypothesis.

Now let H_1 denote the point set composed of K together with the set of all

points $[P]$ of H such that P can be connected with some point of K by a broken line of intervals of length less than ϵ such that the end points of these intervals belong to H. Let H_2 denote the point set composed of all other points of H. H_2 will contain M and it can easily be seen that neither H_1 nor H_2 contains a limit point of the other, since every point of H_2 is at a distance greater than or equal to ϵ from every point of H_1.

LEMMA. *If M is a closed set not disconnecting* a plane S then any two points of $S - M$ can be joined by a simple continuous arc lying in $S - M$.*

Proof. Let P denote any point of $S - M$. Let S_1 denote the point set composed of P together with all points that can be joined to P by a simple continuous arc lying in $S - M$. Let S_2 denote the set $S - M - S_1$, and suppose that S_2 contains at least one point. Now since M does not disconnect S, S_1 contains a limit point of S_2, or S_2 of S_1. Suppose that S_1 contains a limit point P_1 of S_2. Then enclose P_1 within a square K which neither contains not encloses a point of M. This square will enclose a point P_2 of S_2. Then P_1 and P_2 can be joined by a straight line interval lying within K and therefore containing no point of M. Since P_1 can be joined to P by a simple continuous arc lying in $S - M$ it is obvious that P_2 can also. The argument would be similar in the case where S_2 contains a limit point of S_1. Since either leads to a contradiction we have proved that S_2 does not contain even one point.

THEOREM 3. *If M is the sum of a countable number of closed, mutually exclusive point sets $M_1, M_2, M_3, \cdots$, no one of which disconnects a plane S, then M does not disconnect S.*

Proof. Suppose on the contrary that $S - M = S_1 + S_2$, where S_1 and S_2 are mutually exclusive and neither contains a limit point of the other. Let $\overline{M}_1$ denote a point set composed of those points of M that are limit points of S_1 but not of S_2, $\overline{M}_2$ the point set composed of those that are limit points of S_2, but not of S_1, and let $\overline{M}$ denote the point set composed of those points of M that are limit points of neither S_1 nor S_2. There exists in S a countable collection K_{P_1} of squares $K_1, K_2, K_3, \cdots$ obtained in the following manner. Take any point P_1 of S_1 as the center of a square $\overline{K}_1$ of side 2 inches. Let K_1, K_2, K_3, K_4 be the four squares of side one inch each contained in $\overline{K}_1$ and taken in any order. In general let $\overline{K}_n$ be a square of side 2^n inches which has P_1 for its center and has its sides parallel to those of $\overline{K}_1$, and let it be divided into 2^{4n-2} squares, each of side $1/2^{n-1}$ inches, and let these 2^{4n-2} squares follow each other in any order, and let them be the squares $K_{\frac{2^{4n-2}+11}{15}}, \cdots, K_{\frac{2^{4n+2}-4}{15}}$ in the set K_{P_1}.

Now consider the first square K_1' of K_{P_1} which contains P_1 and satisfies

* M is said to disconnect S if $S - M$ is the sum of two mutually exclusive point sets neither of which contains a limit point of the other.

condition (1) that it contain and enclose only points of $S_1 + \overline{M}_1 + \overline{M}$. There evidently exists one such square, since P_1 is not a limit point of S_2. Add to K_1' the first square K_2' of K_{P_1} that encloses no point of the interior of K_1' and that satisfies condition (1) and also condition (2) that it shall contain or enclose at least one point of S_1, and condition (3) that it shall have an interval in common with K_1'. It is evident that there exists a simple closed curve C_2 which is a subset of $K_1' + K_2'$ and such that the interiors of K_1' and K_2' are subsets of the interior of C_2. In general obtain C_3, C_4, C_5, $\cdots$ in the following manner: C_r shall be obtained by adding to C_{r-1} the first square K_r' of K_{P_1} which encloses no point of the interior of C_{r-1} and which satisfies conditions (1) and (2) and contains an interval in common with C_{r-1}. Then C_r is a simple closed curve which is a subset of $C_{r-1} + K_r'$ and whose interior contains the interiors of C_{r-1} and K_r'. It can easily be shown that provided K_r' exists and C_{r-1} can be obtained in this manner, then C_r can also. For let $A_1 A_2$ be an interval common to C_{r-1} and K_r'. Let $A_1 \overline{B} A_2$ (Fig. 2) be an arc

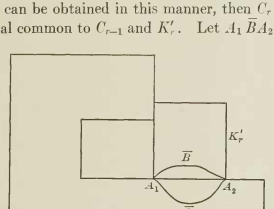

FIG. 2.

lying except for its end points within K_r' and let $A_1 \overline{C} A_2$ be an arc lying except for its end points within C_{r-1}. Then in Theorem 41 of *The foundations of plane analysis situs*,* let J_1 be $C_{r-1} - A_1 A_2 + A_1 B A_2$ and let J_2 be $K_r' - A_1 A_2 + A_1 \overline{C} A_2$. Then J will be the required C_r. Denote by C the set of curves C_1 (K_1') C_2, C_3, $\cdots$.

The sequence C will evidently be infinite unless some C_i contains only points of M. Suppose this to be the case, and suppose first that there is some point P_E of S_1 or S_2 without C_i. The curve C_i divides S into two parts, its interior I, and its exterior E, such that neither of these parts contains a limit point of the other. Now all of C_i must belong to the same set in M,

* See R. L. Moore, *On the foundations of plane analysis situs*, these Transactions, vol. 17 (1916), p. 155.

say M_j, since there does not exist a countable collection of closed, mutually exclusive point sets, consisting of more than one set, whose sum is closed, bounded, and connected.† Then M_j disconnects S; one part of $S - M_j$ is composed of those points of I which do not belong to M_j (P_1 is one such point); the other part is composed of those points of E which do not belong to M_j (by hypothesis there is at least one such point). But this is contrary to the hypothesis of the theorem. Suppose, secondly, that all points of C_i and its exterior belong to M_j. Start from P_1 again to add squares of K_{P_1}, this time adding the extra condition that no square added shall contain or enclose a point of C or its exterior or of M_j; this is possible since M_j and the set composed of C_i together with its exterior are closed sets. Suppose again that some C_k of this new set of curves is a subset of M, and therefore of, say, M_l. If there is some point of S_1 or S_2 without C_k, then, as above, M_l disconnects S, which is contrary to hypothesis. But there must be a point of S_1 or S_2 in the closed, connected, bounded point set made up of $C_i + C_k$ and all points between them, as otherwise this set of points would be the sum of a countable collection of closed, mutually exclusive point sets, subsets of M, and would contain subsets of at least two of the sets $M_1, M_2, M_3, \cdots$, namely of M_j and of M_l. Thus in any case we obtain an infinite sequence C, every curve of which contains at least one point of S_1 and contains only points of $S_1 + \overline{M}_1 + \overline{M}$. Let D_1 denote the sum of the interiors of these curves.

Now suppose there is no point of S_2 without D_1. Then take a point P_2 of S_2 that is within some curve C_{P_2} of C and add squares to it in the manner in which we obtained K_1', K_2', K_3', $\cdots$, except that these squares with their interiors are subsets of $S_2 + \overline{M}_2 + \overline{M}$ and each one contains or encloses at least one point of S_2, and add the extra condition that no one shall contain a point of C_{P_2} or of its exterior. The sum of the interiors of the curves so obtained will be a domain D_2, a subset of the interior of C_{P_2} and therefore not containing all points of S_1. Since D_2 would serve as well for the argument as D_1 we shall suppose that not every point of S_2 is within D_1.

Then since D_1 does not contain all of S it must have some boundary points; let B denote the boundary of D_1. Suppose B contains a point P_1' of S_1. There is a square $\overline{R}$ of the sequence K_{P_1} which encloses or contains P_1' but no point or limit point of S_2 and which encloses a point of a curve of the sequence C. Consider the first curve of C which was obtained by adding a square having an interval in common with $\overline{R}$ or its interior. Let R^* be the square so added. Consider two cases.

Case I. It is given that R^* has an interval in common with $\overline{R}$. If the interiors of these squares are mutually exclusive, R possesses all of the properties necessary for it to be added in obtaining some curve of C, and it will

† Cf. M. Sierpiński, loc. cit.

subsequently be added or enclosed by the addition of some other square of K_{P_0}. If the interiors of R^* and $\bar{R}$ are not mutually exclusive, evidently the interior of $\bar{R}$ must include that of R^*. But $\bar{R}$ precedes R^* in the sequence K_{P_0} and would have been used instead of R^*, since $\bar{R}$ contains an interval of R^* that R^* has in common with that curve of C to which we supposed it added.

Case II. It is given that R^* has an interval in common with the interior of $\bar{R}$. In this case the interior of R^* lies wholly within R but obviously they must also have an interval in common, since R^* was the square added to obtain the first curve of C having an interval in common with $\bar{R}$ or its interior. The argument is the same then as in Case I. We have therefore proved that B contains no point of S_1.

Furthermore B contains no point of S_2. For suppose P_2' is a point of S_2 belonging to B. It is not on any curve of C; therefore it must be a limit point of an infinite number of curves of C. Since P_2' is not a limit point of S_1 there exists a square K_h with P_2' as center, which neither contains nor encloses a point of S_1. Suppose a side of K_h is ϵ_1 inches long. There exists only a finite number of squares of K_{P_0} of side equal to or greater than $\epsilon_1/8$ that have points in common with K_h or its interior. Let Q denote this set of squares. If any square of Q was used in the sequence $K_1', K_2', K_3', \cdots$, let K_n' denote the last square in this sequence that belongs to Q. Then C_n will be the last curve of C formed by adding a square of Q. If no square of Q was so used, let C_n denote any curve of C. Then K_h encloses a point P_1 which lies on no curve of the set $C_1, C_2, \cdots, C_n$, but which does lie on a curve of C following C_n and such that the distance from P_2' to P_1 is less than $\epsilon_1/4$. Then P_1 must be a point on a square of the sequence $K_1', K_2', K_3', \cdots$ of side less than $\epsilon_1/8$ and therefore is at a distance of less than $\epsilon_1/4$ from a point of S_1, since every square of the sequence $K_1', K_2', K_3', \cdots$ contains or encloses a point of S_1. This point of S_1 would then lie within K_h, which leads to a contradiction. We have therefore proved that B is a subset of M.

We can now prove that two closed, mutually exclusive point sets neither of which disconnects S cannot together disconnect S.† For suppose $M = M_1 + M_2$. We have shown above that not every point of S_2 is in D_1 and that no point of S_2 is on B. Let P_0 denote a point of S_2 without D_1. There is a simple continuous arc from P_0 to P_1. Let B_1 denote the first point of B on this arc in the order $P_0 P_1$. Suppose that M_1 is that one of the sets M_1 and M_2 to which B_1 belongs. Since M_1 is a closed set not disconnecting S, P_0 can, by the lemma, be joined to P_1 by an arc not containing any point of M_1. Let B_2 be the first point of B on this arc‡ in the order $P_0 P_1$. Then

† Hausdorff gives a proof for the case when one of the sets is bounded. Cf. F. Hausdorff, *Grundzüge der Mengenlehre*, Leipzig, Veit, 1914, p. 342.

‡ Hereafter in this paper, "arc" and "simple continuous arc" will be considered synonymous terms.

B_2 belongs to M_2. The set $P_0 B_1 + P_0 B_2$ contains as a subset an arc $B_1 B_2$. Let H_1 be a simple closed curve enclosing B_1 but neither containing nor enclosing any point of M_2 (Fig. 3) and containing only one point L_1 of $B_1 B_2$,

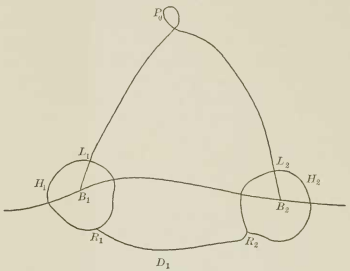

FIG. 3.

and let H_2 be a simple closed curve neither containing nor enclosing any point of H_1 or M_1, but enclosing B_2 and containing only one point L_2 of $B_1 B_2$. Let a point of D_1 within H_1 be joined to a similar point within H_2 by an arc lying within D_1. There will be a subset of this arc, an arc $R_1 R_2$, such that R_1 lies on H_1, R_2 on H_2 and all other points of $R_1 R_2$ lie without both H_1 and H_2. Then there is a simple closed curve J_1 composed of $L_1 L_2 + R_1 R_2$ together with either arc $L_1 R_1$ on H_1 and either arc $L_2 R_2$ on H_2. It is evident that the points of B on or within J_1 that belong to M_1 can be enclosed in a finite number of circles no one of which contains or encloses a point of $L_1 L_2 + L_2 R_2 + R_1 R_2 + M_2$. And similarly those points of B on or within J_1 that belong to M_2 can be enclosed in a finite number of circles no one of which contains or encloses a point of $L_1 L_2 + L_1 R_1 + R_1 R_2 + M_1$ or a point on or within a circle of the first set. Then, clearly, a point on $L_1 L_2$, and therefore without D_1, could be joined to a point on $R_1 R_2$, and therefore within D_1, by an arc lying within J_1 and without both these sets of circles, and therefore not containing a point of B. Since this leads to a contradiction we have shown that if M consists of only two sets it cannot disconnect S. This result can evidently be extended to the case where M consists of any finite number of sets.

Consider the arcs $P_0 B_1$ and $P_0 B_2$ above. Let A_2' denote the first point of $B_2 P_0$ on $P_0 B_1$. If A_2' is different from P_0 it is evident that a point on $B_2 A_2'$ very near A_2' can be joined to P_0 by an arc lying without D_1 and con-

taining only P_0 in common with $P_0 B_1$. From this together with $B_2 A_2'$ we obtain an arc $P_0 B_2$ which lies, except for B_2, without D_1 and has only P_0 in common with $P_0 B_1$. Let M_{B_1} denote that set of M to which B_1 belongs, M_{B_2} the set to which B_2 belongs. Since $M_{B_1} + M_{B_2}$ does not disconnect S, P_0 can be joined to P_1 by an arc not containing any point of M_{B_1} or M_{B_2}. Let B_3 denote the first point of B on this arc in the order $P_0 P_1$, and let M_{B_3} denote that set of the collection M to which B_3 belongs. As above, let A_3' denote the first point $B_3 P_0$ has on the sum of the arcs $P_0 B_1$ and $P_0 B_2$. By taking a point on $B_3 A_3'$ very near A_3', and drawing a suitable arc to P_0, we obtain an arc $P_0 B_3$ which has only P_0 in common with $P_0 B_1$ or $P_0 B_2$, and which lies, except for B_3, without D_1. Similarly obtain the arc $P_0 B_4$ where B_4 belongs to B and to M_{B_4}, such that $P_0 B_4$ has only P_0 in common with $P_0 B_1$, $P_0 B_2$, or $P_0 B_3$, and lies, except for B_4, without D_1. Now it is evident that the sum of two of these arcs is an arc crossing the sum of the other two. Suppose for instance that the arc $P_0 B_1 + P_0 B_3$ crosses the arc $P_0 B_2 + P_0 B_4$. Let H_3 be a simple closed curve (Fig. 4) enclosing B_1 but

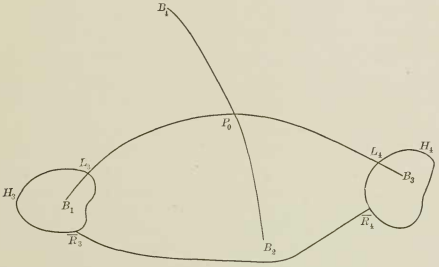

FIG. 4.

neither containing nor enclosing any point of $P_0 B_2$, $P_0 B_3$, $P_0 B_4$, M_{B_2}, M_{B_3}, or M_{B_4} and such that it contains only one point L_3 of the arc $P_0 B_1$. Let H_4 be a simple closed curve enclosing B_3, containing only one point L_4 of the arc $P_0 B_3$, and neither containing nor enclosing any point of H_3, $P_0 B_1$, $P_0 B_2$, $P_0 B_4$, M_{B_1}, M_{B_2}, or M_{B_4}. Now consider the first curve of C, C_A, which has points within both H_3 and H_4. There is an arc $R_3 R_4$, a subset of C_A, such that R_3 is on H_3, R_4 is on H_4 and all other points of $R_3 R_4$ lie without both H_3 and H_4. Select one of the arcs $L_3 R_3$ on H_3 and one of the arcs $L_4 R_4$ on H_4 and let those selected be denoted by $L_3 R_3$ and $L_4 R_4$ throughout

the discussion. Then $L_3 L_4 + L_4 R_4 + R_3 R_4 + L_3 R_3$ is a simple closed curve J_2. Now let $\overline{R}_3 \overline{R}_4$ be an arc on C_A lying on J_2 or within J_2 except for its end points, and having $\overline{R}_3$ on H_3 and $\overline{R}_4$ on H_4 and all other points without both H_3 and H_4, such that if we consider the simple closed curve $\overline{J}_2$ which $\overline{R}_3 \overline{R}_4$ forms with that arc of J_2 (from R_3 to $\overline{R}_4$) that contains $L_3 L_4$, there is no arc of C_A (except $\overline{R}_3 \overline{R}_4$) lying on $\overline{J}_2$ or within $\overline{J}_2$ except for its end points, and having one point on H_3 and one on H_4. It is possible that the arc $\overline{R}_3 \overline{R}_4$ will be the arc $R_3 R_4$. Now it is evident that either $\underset{\frown}{P_0 B_2}$* or $\underset{\frown}{P_0 B_4}$ will lie within $\overline{J}_2$. Suppose $\underset{\frown}{P_0 B_2}$ does. Then B_2 will lie on or within $\overline{J}_2$. The point B_2 is a limit point of an infinite sequence of curves of C following C_A. Let $\overline{P}_1, \overline{P}_2, \overline{P}_3, \cdots$ denote a sequence of points on successive curves of this sequence such that B_2 is a sequential limit point of the set $\overline{P}_1, \overline{P}_2, \overline{P}_3, \cdots$. If B_2 is on $\overline{J}_2$ these points may all coincide with B_2; if not, let them be chosen so that they lie within $\overline{J}_2$.

Denote by $\overline{C}_i$ that curve of C on which $\overline{P}_i$ lies, and let $\overline{N}_i$ be the last point of $\overline{C}_i$ starting from $\overline{P}_i$ in either order around $\overline{C}_i$ such that $\overline{P}_i \overline{N}_i$ lies on or within $\overline{J}_2$. It will be seen that every $\overline{N}_i$ will have to be on $L_3 \overline{R}_3$ or $L_4 \overline{R}_4$, for if any $\overline{N}_i$ is on $\underset{\frown}{\overline{R}_3 \overline{R}_4}$ there must be a point F of $\overline{C}_i$ very near $\overline{N}_i$, without $\overline{J}_2$ and therefore on the opposite side of C_A from $\underset{\frown}{P_0 B_2}$; for if B_2 is not on C_A it can be joined, because of the condition put upon $\overline{R}_3 \overline{R}_4$, to any point P on $\overline{R}_3 \overline{R}_4$ by an arc having only the point P on C_A. Now since $\underset{\frown}{P_0 B_2}$ is without C_A, F must be within C_A. This is impossible, since F is on a curve following C_A in C, and is therefore on or without C_A. Now the set of arcs $\overline{P}_i \overline{N}_i$ (where $i = 1, 2, 3, \cdots$) determines a limit set Y such that Y is a closed, connected set, every point of which is on or within $\overline{J}_2$. Since B_2 belongs to Y, Y must be a subset of M_{B_2}. But there will be a point of Y on either $L_3 \overline{R}_3$ or $L_4 \overline{R}_4$, namely a limit point of the set of points $\overline{N}_1, \overline{N}_2, \overline{N}_3, \cdots$. This leads to a contradiction, for neither H_3 nor H_4 contains a point of M_{B_2}. The supposition that M disconnects S is therefore proved false.

THEOREM 4. *If M_1 and M_2 are two closed, connected, bounded point sets, neither of which disconnects a plane S, a necessary and sufficient condition that their sum, M, shall disconnect S is that $\overline{M}$, the set of points common to M_1 and M_2, be not connected.*

Proof. The condition is necessary. For suppose that $\overline{M}$ is connected; we can prove that M does not disconnect S. For suppose $S - M = S_1 + S_2$ where S_1 contains no limit point of S_2 nor S_2 of S_1. Let $\overline{S}$ be a square enclosing M. Then it is evident that either S_1 or S_2 must lie within $\overline{S}$. Suppose S_1 does. Let P_1 be a point of S_1, and consider all points which lie

* If $P_0 B_2$ is an arc, $\underset{\frown}{P_0 B_2}$ denotes the point set $P_0 B_2 - P_0 - B_2$.

with P_1 in a connected subset of S_1. It will be seen that since M is closed, these points form a domain D_s, a subset of S_1, and of the interior of $\overline{S}$, such that the boundary of D_s is a subset of M. Let M' denote the point set $M_1 - \overline{M}_2$ and M'' the point set $M_2 - \overline{M}$. Let P_2 denote a point of S_2 that is without $\overline{S}$; by the lemma, P_2 can be joined to P_1 by an arc not containing any point of M_2. Let P' denote the first point the arc $P_2 P_1$ has on

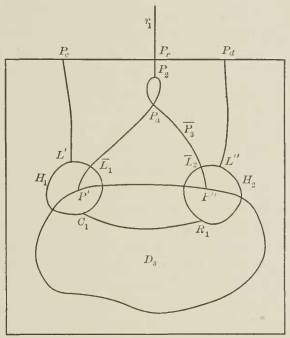

FIG. 5.

the boundary of D_s (Fig. 5), let P_2 be joined to P_1 by an arc not containing any point of M_1, and let P'' denote the first point of this arc $P_2 P_1$ on the boundary of D_s. Then P' belongs to M' and P'' to M''. Then there exists an arc $P' \overline{P}_3 P''$, a subset of $P_2 P' + P_2 P''$, which lies, except for its end points, without D_s. As in Theorem 3, there exists a simple closed curve H_1 enclosing P', containing only one point $\overline{L}_1$ of the arc $P' \overline{P}_3 P''$, and neither containing nor enclosing a point of M_2, and there exists a corresponding curve H_2 enclosing P'', containing only one point $\overline{L}_2$ of the arc $P' \overline{P}_3 P''$ and neither containing nor enclosing a point of M_1 or H_1, and there exists an arc $C_1 R_1$ in D_s having only C_1 on H_1 and R_1 on H_2. Then there is a simple closed curve $\overline{J}_1$, composed of $L_1 \overline{P}_3 \overline{L}_2 + C_1 R_1$ together with the arcs $\overline{L}_1 C_1$ on H_1 and $\overline{L}_2 R_1$ on H_2 so chosen that $\overline{J}_1$ encloses P' and P''.

We can show as in Theorem 3 that if those points of the boundary of D_s that this curve contains or encloses belong to $M' + M''$ then two points, one without D_s and the other within D_s, can be joined by an arc lying within

$\bar{J}_1$ and containing no point of the boundary of D_s. Since this is impossible $\bar{J}_1$ must enclose some point of $\bar{M}$, and since $\bar{M}$ is connected and since no point of $\bar{M}$ is on $\bar{J}_1$, $\bar{M}$ must lie wholly within $\bar{J}_1$.

A ray of a straight line can be drawn from P_2, lying wholly without $\bar{S}$, and there exists a ray r_1, a subset of the first ray together with one of the arcs $P_2 P_1$, such that r_1 lies without $\bar{J}_1$ except for its end point P_a on $P' P_3 P''$. Let $\bar{S}'$ denote a square enclosing $\bar{S}$ and both arcs $P_1 P_2$ and containing only one point P_r of r_1. If L' is a point on H_1 and $\bar{J}_1$ sufficiently near $\bar{L}_1$ there exists an arc $P_c L'$ such that P_c is a point on $\bar{S}'$ and such that $P_c L'$ lies without D_s, $P_c L' - L'$ lies without $\bar{J}_1$ and $P_c L' - P_c$ lies within $\bar{S}'$. Similarly a point L'' of H_2 and $\bar{J}_1$ can be joined by an arc to the point P_D of $\bar{S}'$, and in such a way that the arc $P_D L''$ contains no point of the arc $P_c L'$. Let $L' L''$ denote the arc composed of $P_c L' + P_D L''$ together with that arc of $\bar{S}'$, from P_c to P_D, which does not contain P_r. The arc $L' L''$ together with the arc $L' C_1 R_1 L''$ of $\bar{J}_1$ makes a simple closed curve $\bar{J}_2$ whose interior has no point in common with that of $\bar{J}_1$. Since $\bar{J}_2$ encloses points of D_s we can show as above that it must contain all of $\bar{M}$. Since this is obviously impossible, the supposition that $\bar{M}$ was connected is proved to be false.

The condition is sufficient. For suppsoe M is not connected. Then it is the sum of two closed, mutually exclusive point sets, $\bar{M}_1$ and $\bar{M}_2$. We can show that in this case $S - M = S_1 + S_2$ where S_1 and S_2 are mutually exclusive and neither contains a limit point of the other.

The point set $\bar{M}_1$ contains some point $\bar{P}_1$ which is a limit point of M'' since M_2 is connected and $\bar{M}_1$ and $\bar{M}_2$ are closed point sets. Let J_{p_1} denote a circle enclosing $\bar{P}_1$ but neither containing nor enclosing any point of $\bar{M}_2$. This circle contains a point P_1^* of $M_2 - \bar{M}_1 - \bar{M}_2$, and there is a ray of an open curve from P_1^* not containing any point of M_1, since M_1 is closed and bounded and does not disconnect S. And a subset of this ray is a ray r_2, having an end point on J_{p_1} and lying, except for this end point, without J_{p_1}. Now, by the Heine-Borel theorem, $\bar{M}_2$ can be enclosed in a finite set of circles no one of which contains or encloses any point of $\bar{M}_1$ or J_{p_1} or r_2. Consider any one of these circles together with all of the set that are connected with it. By Theorem 42 of *The foundations*† there exists a simple closed curve J_{c_2}, a subset of these circles, whose interior contains all of their interiors. It will not contain any point of $\bar{M}_1$, but will enclose some points of $\bar{M}_1$; let M_2^* denote the set of these points. The curve J_{p_1} with r_2 will be without J_{c_2}. Now, as above, there exists a ray r_3 which has its end point on J_{c_2}, lies, except for this point, without J_{c_2} and contains no point of M_1. Then all points of $\bar{M}$ that are without J_{c_2} can be enclosed in a finite set of circles no one of which contains or encloses any point of J_{c_2} or r_3. If these circles do not form a connected

† Cf. R. L. Moore, loc. cit., p. 156.

point set they can be joined by a finite number of arcs not containing any point of $J_{c_2} + r_3$. For J_{c_2} with its interior does not disconnect S, and r_3 does not disconnect S and they have only one point in common; therefore, by the first part of this theorem, J_{c_2} with its interior and r_3 does not disconnect S. These arcs can be covered by a finite set of circles not containing any point of J_{c_2} or r_3, and hence from all these circles a simple closed curve J_{c_1} can be obtained which has the following properties: it is wholly without J_{c_2} and does not enclose it; it encloses all points M_1^*, of $\overline{M}$, that J_{c_2} does not enclose; therefore it contains no point of $\overline{M}$.

The points of M_2 that are on J_{c_1} or J_{c_2} or without both J_{c_1} and J_{c_2} form a closed set, a subset of M''. This set can be covered by a finite set of circles, T, such that no circle of T contains or encloses any point of M_1. Now there is a ray of an open curve, not containing any point of M_1, from some point within each of these circles. Let $[r_t]$ denote the sum of such rays. And let M_1 be covered by a finite set of circles no one of which contains or encloses any point of the circles T, or of the point set $[r_t] + r_2 + r_3$. Since M_1 is connected this set of circles will be connected, and, as above, there exists a simple closed curve J_1 which is a subset of them and encloses all of their interiors and therefore all of M_1. The curve J_1 cannot wholly enclose either J_{c_1} or J_{c_2} since it contains no point of r_2 or r_3.

Let S^* be a square enclosing J_{c_1}, J_{c_2}, J_1 and M_2. It is evident that there is an arc $P_3 P_4$ composed of a finite number of straight line intervals lying within S^*, except for the points P_3 and P_4 which are on S^*, and which separates the interior of S^* into two parts such that J_{c_1} lies wholly within one part and J_{c_2} within the other. Let A be a point of $\overline{M}$ within J_{c_1} and B a point of $\overline{M}$ within J_{c_2}. By Theorem 1 of this paper there is a connected subset of M_2 lying within J_1 such that A and some point A' of J_1 are limit points of it. Moreover this set will lie within J_{c_1} since it is connected, A lies within J_{c_1} and J_1 encloses no point of M_2 that is on J_{c_1}. Similarly there exists a connected subset of M_2 lying within J_1 and J_{c_2} such that B and some point B' on J_1 are limit points of it.

Now there is a subset, CD, of $P_3 P_4$ which satisfies the following conditions: it is an arc lying except for C and D within J_1; C and D are on J_1 and separate A' and B' on J_1. For consider all arcs of $P_3 P_4$ which lie except for their end points within J_1 and have their end points on J_1. There are a finite number of these, since J_1 is composed of a finite number of arcs of circles and $P_3 P_4$ of a finite number of straight line intervals. Let $L_1, L_2, L_3, \cdots$, denote the set of all such arcs and let L denote the point set $L_1 + L_2 + L_3 \cdots$. Suppose that no L_i separates A' and B' on J_1. Denote by a' and a'' the two arcs $A'B'$ on J_1. If any arc L_i of L has an end point on a' for instance let X_1 denote the first such end point on a' in the order $A'B'$ and X_2 the other

end point of this arc. Then consider the simple closed curve $X_1 A' + a'' + B' X_2 + X_2 X_1$. It will be seen that A' and B' can be joined by an arc lying on or within this curve which contains no point of L_i. By a finite number of repetitions of this process it is evident that we can obtain a simple closed curve C_x, a subset of J_1 plus its interior, containing A' and B' but enclosing no point of L. A' and B' could then be joined by an arc lying on or within C_x and containing no point of L and hence no point of $P_3 P_4$. Since this is impossible there must exist an arc CD, having the properties stated above.

Since $\overset{\frown}{CD}$, being a subset of $P_3 P_4$, lies without both J_{c_1} and J_{c_2} and since it lies within J_1, it contains no point of M_2. Therefore of the two regions into which it divides the interior of J_1,* one must contain A, the other B. Let R denote the interior of J_1. Let $F_1 F_2 F_3$ be an arc as indicated in Fig. 6

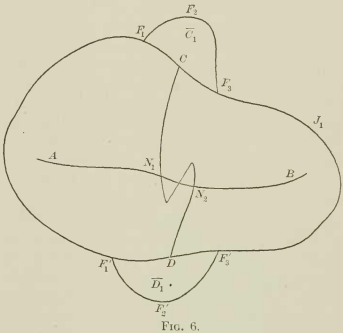

Fig. 6.

with its end points on J_1 and lying, except for its end points, without J_1 and wholly without J_{c_1} and J_{c_2} and such that the simple closed curve $F_1 C F_3 F_2 F_1$ neither contains nor encloses a point of M_1 or M_2. Let $\overline{C}_1$ be a point within this curve. Let $F_1' F_2' F_3'$ be a similar arc forming part of a simple closed curve $F_1' D F_3' F_2' F_1'$ and let $\overline{D}_1$ be a point within it. Now if we suppose that M does not disconnect S, let N_1 be the first point of the arc CD on M_1. Then by the first part of this theorem, $CN_1 + M$ does not disconnect S. Let N_2 be the first point the arc DC has on M_1 (N_1 and N_2 must exist since M_1 is con-

* Cf. R. L. Moore, loc. cit., p. 141.

nected and CD divides R into two parts, of which one contains A, and the other contains B). By several applications of the first part of this theorem, we can show that $M + CN_1 + DN_2 + F_1CF_3 + F_1'DF_3'$ does not disconnect S. Let H denote this point set. Then $\overline{C}_1$ and $\overline{D}_1$ can be joined by a broken line containing no point of H. By extending this broken line we get a broken line CED which has only C and D on H. By Theorem 43 of *The foundations*† there is a simple closed curve $\overline{C}$, a subset of J_1 together with the broken line CED whose interior is a subset of R, which encloses N_1 but no point of the arc CED. Then $\overline{C}$ will enclose all of $M_1 + (CN_1 - C) + (DN_2 - D)$, since this is a connected point set and has no point on $\overline{C}$. The curve $\overline{C}$ will contain F_1CF_3, $F_1'DF_3'$, A' and B' since it encloses A and B and contains no point of the previously described connected sets of M_2 between A and A', and B and B'. Then it is evident that there is a curve $\overline{C}'$, which has all the above mentioned properties of $\overline{C}$ except that it contains only the points C and D of the broken line CED and which is such that (1) it is a subset of $\overline{C}$ plus its interior, (2) those points of $\overline{C}'$ that are not on $\overline{C}$ do not belong to M.

It is possible that the broken line segment $CD - C - D$ (a subset of P_3P_4) will not lie within $\overline{C}'$. Consider the arcs of the broken line CD which lie except for their end points without $\overline{C}'$ and have their end points on $\overline{C}'$. Let $C'D'$ denote such an arc. C' and D' are points of R. Let E' denote a point of $C'D'$ without $\overline{C}'$. The points C' and D' can be joined by an arc $C'F'D'$ such that $C'F'D'$ lies within $\overline{C}'$ and therefore within J_1. Then the simple closed curve $C'E'D'F'C'$ is a subset of R, and therefore the arc $C'GD'$ of $\overline{C}'$ that it encloses must be a subset of R, and therefore it must belong to that part of $\overline{C}'$ which does not belong to $\overline{C}$ and which therefore contains no points of M. Now there are only a finite number of arcs such as $C'E'D'$, that are parts of the broken line CD, which lie except for their end points without $\overline{C}'$ and have their end points on $\overline{C}'$, and we see that each one can be replaced by an arc lying on $\overline{C}'$ and containing no point of M_2. This gives us a continuous curve from C to D lying on or within $\overline{C}'$ and containing no point of M_2, and there is an arc, a subset of this curve, having the same properties. It is evident that this arc can be replaced by an arc CE^*D lying within $\overline{C}'$ except for C and D and containing no point of M_2. Now C and D separate A' and B' on $\overline{C}'$. For if they did not, A' and B' could be joined by an arc a_1 lying, except for A' and B', within $\overline{C}'$, and C and D could be joined by a similar arc a_2 such that a_2 had no point in common with a_1. But both these arcs would lie except for their end points within $\overline{C}$ and J_1 and have their end points on $\overline{C}$ and J_1. This leads to a contradiction, since C and D separate A' and B' on J_1.

Now the arc CED, together with the arc CE^*D, forms a simple closed

† See R. L. Moore, loc. cit., p. 157.

curve J_3 containing no point of M_2, but such that one of the points A and B is within it, and the other without it. This is impossible since M_2 is a connected point set. Therefore the supposition that $\overline{C}_1$ and $\overline{D}_1$ can be joined by an arc containing no point of H is false and since H disconnects S, M must disconnect S. Evidently one of the two sets into which M separates S is the point set, S_1, composed of $\overline{C}_1$ together with all points that can be joined to it by arcs containing no point of M. The set S_2 will then be the point set $S - M - S_1$.

THEOREM 5.* *If M_1 and M_2 are two closed, bounded, connected point sets in a plane S, such that neither M_1 nor M_2 disconnects S and such that M_1 and M_2 have in common only K_1 and K_2, where K_1 and K_2 are mutually exclusive connected sets, then $S - M_1 - M_2$ is the sum of just two mutually exclusive, connected domains.*

Proof. We have shown above that under the conditions of this theorem $S - M_1 - M_2$ is not connected. Suppose then that it is the sum of more than two mutually exclusive connected domains. There will exist three points, P_1, P_2 and P_3, no two of which can be joined by an arc containing no point of $M_1 + M_2$. It is evident that in the preceding theorem the curve J_1 could have been constructed in such a way that P_1, P_2 and P_3 were without it; for there exist three open curves, containing P_1, P_2 and P_3 but containing no point of M_1, and J_1 could have been drawn so as not to contain any point of these open curves. We shall suppose that J_1 has been so drawn. We can furthermore suppose that J_1 is replaced by a polygon W, satisfying the conditions which J_1 satisfies.

Now P_1 and P_2 can be joined by an arc made up of a finite number of straight line intervals containing no point of M_2, since M_2 does not disconnect S; and since $M_2 + P_1 P_2$ does not disconnect S, it is obvious that P_2 and P_3 can be joined by a similar arc which contains no point of M_2 and has only P_2 in common with $P_1 P_2$. Similarly there is a broken line, an arc $P_3 P_1$, containing no point of M_2, and having only P_3 and P_1 in common with $P_1 P_2 + P_2 P_3$. Now any one of these three arcs will contain only a finite number of intervals lying except for their end points in W and having their end points on W. These intervals will be of two kinds, those whose end points separate A' and B' on W, and those whose end points do not separate A' and B' on W. Consider an interval $\overline{X}_1 \overline{X}_2$ of the second sort. Let $\overline{X}_1 X' \overline{X}_2$ denote that arc on W which does not contain A' or B'. Then obviously $\overline{X}_1 \overline{X}_2 + \overline{X}_1 X' \overline{X}_2$ is a simple closed curve enclosing no point of K_1 or K_2, and by the method employed in Theorem 3, we can draw an arc $\overline{X}_1 \overline{X}_2$ such that $\overline{X}_1 \overline{X}_2$ is within this curve and contains no point of either

* Rosenthal gives a proof for the case in which each of the sets K_1 and K_2 reduces to a single point. See A. Rosenthal, loc. cit., p. 102, Theorem 6.

M_1 or M_2. Let the original interval $\overline{X}_1\,\overline{X}_2$ be replaced by this arc $\overline{X}_1\,\overline{X}_2$ and let this process be carried out for every such interval, of the second kind, on each of the three arcs. Now each one of the three must have at least one interval of the first kind; for since no one of these arcs has any point of M_2 on it and all of M_1 lies within W and every interval of the second sort of the arcs lying within W has been replaced by an arc containing no point of either M_1 or M_2, if one of these arcs, $P_1\,P_2$, for instance, had no interval of the first sort on it, it would be replaced by an arc $P_1\,P_2$ which had no point of M_1 or M_2 on it, which is contrary to our supposition. Let $Y_1\,Y_2$ be the first interval on $P_1\,P_2$ in the order $P_1\,P_2$ which separates A' and B' on W, let $Y_3\,Y_4$ be a similar arc on $P_2\,P_3$, and let $Y_5\,Y_6$ be a similar arc on $P_3\,P_1$. Some two of the points Y_1, Y_3, Y_5 must lie on the same arc $A'\,B'$ of W; suppose Y_1 and Y_3 do (Fig. 7). Now consider the simple closed curve

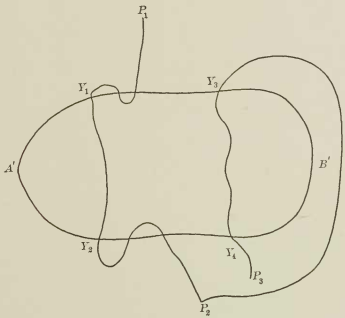

FIG. 7.

$\overline{Y} = Y_1\,Y_3\,Y_4\,Y_2\,Y_1$. Since it contains neither A' nor B' it encloses no point of either K_1 or K_2. Now there is no connected subset of M_2 lying on or within $\overline{Y}$ and having a point $\overline{X}_3$ on $Y_1\,Y_3$ and a point $\overline{X}_4$ on $Y_2\,Y_4$. For such a set would separate A' and B' on W and would divide the interior of W into two sets of which one contains K_1 and the other contains K_2. But K_1 and K_2 are connected within W by a set of points belonging to M_1. This set would have to cross the subset of M_2 lying within Y. But this is obviously impossible. Then by Theorem 2 there is a division of the points of M_2 that lie on or within $\overline{Y}$ into two mutually exclusive closed sets, Z_1 and Z_2, such that Z_1

contains all the points of M_2 that lie on $Y_1 Y_3$. Now all points of Z_1 can be enclosed in a finite number of circles not containing any point of $Y_3 Y_4 + Y_4 Y_2 + Y_2 Y_1$ or any point of Z_2 or M_1, and evidently there is an arc lying on or within $\overline{Y}$, a subset of these circles together with intervals of $Y_1 Y_3$, which contains no point of either M_1 or M_2. Then P_1 and P_2 are joined by an arc not containing any point of either M_1 or M_2. Since this is contrary to our supposition the theorem is proved.

UNIVERSITY OF TEXAS,
AUSTIN, TEX.

ON THE LOCATION OF THE ROOTS OF CERTAIN TYPES OF POLYNOMIALS*

BY

J. L. WALSH

When we study the dependence of a variable on k other variables which vary independently, our problem may be very much simplified if we can consider all or some of these independent variables to coincide and thus study the dependence of our original dependent variable on one new variable or at least on a number of new independent variables less than k. The present writer has recently published a theorem (Theorem II, below) which enables us to make a reduction of this sort in the study of the relations between the roots of certain types of polynomials. The present paper aims to prove Theorem I (below), which is a much more general result of the same nature, and to indicate various applications of that theorem. The applications given are extremely simple and follow from Theorem I with practically no further machinery.† The most interesting application is Theorem VI.

Our problem is, more explicitly, to study the geometric relationship of the roots of a polynomial

$$f(z) = (z - a_1)(z - a_2) \cdots (z - a_m)$$

to the roots of a related polynomial

$$\phi(z) = (z - b_1)(z - b_2) \cdots (z - b_n)$$

* Presented to the Society, December 29, 1920, and April 23, 1921.

† When this paper was first offered for publication, the writer believed Theorem I to be new. Professor D. R. Curtiss has kindly pointed out its connection with a theorem due to Grace and has indicated an entirely new proof of Theorem I; the reader will refer to Professor Curtiss's note which immediately follows this paper.

The *point of view* of the present paper in the proof and application of Theorem I seems to be new, and also the results obtained except where otherwise stated.

This paper is the development of a short note published in the Paris Comptes Rendus, March, 1921, to which explicit reference is made later, and which contained in outline the proof of Theorem I. In the interval between the publication of that note and the publication of the present paper, there have appeared a number of other papers dealing with Grace's Theorem. See Szegö, Mathematische Zeitschrift, vol. 13 (1922), pp. 28–55; Cohn, Mathematische Zeitschrift, vol. 14 (1922), pp. 110–148; Egerváry, Acta Litterarum ac Scientiarum, Regiae Universitatis Hungaricae Francisco-Josephinae, vol. 1 (1922), pp. 39–45; Fekete, same Journal, vol. 1 (1923), pp. 98–100.

The present paper has very little in common with these other papers, but with Szegö's Theorem 9 the reader should compare our Theorem X, and with Szegö's Theorems 13–15 compare our Theorem IV and also Walsh, American Mathematical Monthly, vol. 29 (1922), pp. 112–114.

whose roots are supposed to be known.* When we study the functional dependence of a root of f, say a_1 for definiteness, on the b_i, it may occur that without changing a_1 we can replace the n roots b_i by n roots of a polynomial ϕ which coincide at some point b which bears a simple relation to the roots b_i. Then we can study the dependence of a_1 on b instead of on $b_1, b_2, \cdots, b_n$, which is frequently a simplification of our problem.

To the plane of the complex variable we adjoin the point at infinity; infinity is to be considered simply as an ordinary value of the variable. As a geometrical consequence we shall consider the term *circle* to include the possibility of a straight line. We shall have occasion to deal with *circular regions*, by which we mean a closed portion of the plane bounded by a single circle, that is, the interior of a circle, the exterior of a circle (including the point of infinity), a half-plane, a point, or the entire plane; the points of the boundary are always to be included in the region.

I. A general theorem

We proceed to the proof of our main result:

THEOREM I. *Let $f(z)$ be a polynomial in z whose coefficients are polynomials linear in and symmetric in each of the sets of variables*

$$\{\alpha_1, \alpha_2, \cdots, \alpha_k\}, \quad \{\beta_1, \beta_2, \cdots, \beta_l\}, \quad \cdots, \quad \{\lambda_1, \lambda_2, \cdots, \lambda_q\}.\dagger$$

Let these points $\{\alpha_i\}, \{\beta_i\}, \cdots, \{\lambda_i\}$ lie in circular regions $C_\alpha, C_\beta, \cdots, C_\lambda$. Then for any fixed values of these variables and of z we can always make all the $\{\alpha_i\}$ coincide in C_α, all the $\{\beta_i\}$ coincide in C_β, etc., without altering the value of $f(z)$.

The theorem also obtains if we replace $f(z)$ by the quotient of two polynomials of the type described, except that we are to consider the conclusion of the theorem satisfied if these two polynomials vanish simultaneously.‡ Of course if we have two polynomials of the type required, their sum is also of that type.

We shall prove the theorem considering the $\{\beta_i\}, \{\gamma_i\}, \cdots, \{\lambda_i\}$ always as fixed and showing that we can make the $\{\alpha_i\}$ coincide in C_α as stated. Then we can consider the $\{\alpha_i\}$ fixed and coincident, and the $\{\gamma_i\}, \cdots, \{\lambda_i\}$ fixed; our former reasoning will show that the $\{\beta_i\}$ can be made to coincide. Continued reasoning in this manner will evidently complete the proof of the theorem. It will be convenient to assume that the value of $f(z)$ considered

* Professor Curtiss has recently published a very interesting report on this general field, Science, vol. 55 (1922), pp. 189–194.

† The coefficients of $f(z)$ need not be homogeneous in each of these sets of variables, but each coefficient must be a linear combination of the elementary symmetric functions of each of these sets with coefficients linear combinations of the elementary symmetric functions of the other sets. These linear combinations may, moreover, contain constant terms.

‡ This is what actually occurs in the situation of Theorem II if we choose P inside C.

in the hypothesis is zero; this involves no loss of generality, for the addition of a constant term to $f(z)$ does not alter the properties required.

We now wish to show that the α_i can all be made to coincide. If the region C_α is a point, the statement is trivial. If we prove the theorem where C_α is a circular region not the whole plane we have proved it where C_α is the whole plane.

Consider the α_i to vary independently and to have the region C_α as their common locus. Then the relation

$$f(z) = 0 \tag{1}$$

defines z as an analytic function of the α_i and hence z will have a certain locus Z; this locus Z will be a closed point set since C_α is closed. If (1) degenerates and does not effectively contain z at all, we introduce an auxiliary variable ζ by placing

$$F(\zeta) \equiv f(z) - \zeta = 0;$$

the new function $F(\zeta)$ surely contains ζ and we may reason with it as before for $f(z)$.

Let α be any point interior to C_α. There is a certain locus Z of points z corresponding to the null-circle α as the locus of the $\{\alpha_i\}$. Make use of the auxiliary circle C of the coaxial family determined by α and C_α. Let C commence with the position α and gradually enlarge and coincide with C_α, the region bounded by C and containing α always considered as the locus of the $\{\alpha_i\}$. The locus Z also grows larger and varies continuously with C (except as noted below), for the roots z of (1) are continuous functions of the $\{\alpha_i\}$. When the region C continues to enlarge, beyond C_α if necessary, the locus Z eventually becomes the entire plane; this happens ordinarily before C has swept through the entire plane. The point z with which we start is a point of the locus Z surely when C coincides with C_α and possibly before.

The case that the locus Z does not vary continuously with C occurs only if, for some choice of the $\{\alpha_i\}$, (1) vanishes identically in z. When this occurs, the roots $\{z\}$ of (1) do not vary continuously with the $\{\alpha_i\}$. The locus Z may enlarge suddenly and become the entire plane while its boundary does not necessarily sweep over the whole plane.

Whether or not this phenomenon takes place, corresponding to our original point z there is some circle C such that for no circle C', the region C' smaller than and entirely contained in the region C, can z be a point of the locus Z of the roots of (1). This statement and in fact Theorem I as well are true if z is a point of the locus corresponding to α. The statement is true in any other case. For if there is a sequence of sets of points $\{\alpha_i\}$ in a sequence of regions each contained in the preceding, there is at least one limiting set of points $\{\alpha_i\}$ contained in all the regions and hence in the limit region. The relation (1) obtains for this limit set.

Our entire proof of Theorem I rests on the remark that when we fix z and all but two of the points $\{\alpha_i\}$, equation (1) becomes a homographic and involutory relation between the other two of these points and hence when one of these points traces a circle the other also traces a circle. The proof is complicated, however, by the possibility that this relation may degenerate and may effectively contain but one or neither of these variables.*

We prove now a very special case of Theorem I, namely, that if a circular region C contains two points α_1 and α_2 connected by a relation of the form

$$a_1\alpha_1\alpha_2 + a_2(\alpha_1 + \alpha_2) + a_3 = 0, \tag{2}$$

then C contains a root α' of the equation

$$a_1\alpha'^2 + 2a_2\alpha' + a_3 = 0. \tag{3}$$

Transform (2) by an auxiliary linear transformation so that the point at infinity is a double point of the transformation (α_1, α_2) defined by (2). The line through α_1 and α_2 is transformed into itself by the transformation (α_1, α_2), for the three points $(\alpha_1, \alpha_2, \infty)$ go into $(\alpha_2, \alpha_1, \infty)$. When α_1 is moved on this line, α_2 moves on the line in the opposite sense, for if the two points move always in the same sense, when α_1 is moved along the finite segment from its original position to the original position of α_2, α_2 has moved in the same sense, has at no time coincided with the point at infinity and hence has not reached the original position of α_1. The points α_1 and α_2 moving on the line joining them in opposite senses must have a double point D between them. It is thus evident that the region C which contains α_1 and α_2 must contain D or the point at infinity and hence a solution of (3).

The reasoning just used supposes implicitly that the relation (2) does not degenerate. If (2) does not contain α_2 for a particular value of α_1, that value of α_1 is a solution of (3). If (2) is satisfied identically, any values α_1 and α_2 suffice.

We now return to the proof of Theorem I in its general form. Let C be the

* Thus we may start with the auxiliary polynomial

$$F(z) = (z - \alpha_1)(z - \alpha_2)\cdots(z - \alpha_k) \qquad k > 3$$

and consider for the polynomial of Theorem I

$$f(z) = F'(z).$$

The relation

$$f(z) = 0 \tag{1'}$$

if $\alpha_{k-1} = \alpha_k = z$ contains effectively none of the variables $\alpha_1, \alpha_2, \cdots, \alpha_{k-2}$. If $\alpha_k = z$, we have (1') reducing to

$$(z - \alpha_1)(z - \alpha_2)\cdots(z - \alpha_{k-1}) = 0,$$

and hence we have no effective homographic relationship between α_{k-1} and α_k.

This example is not an unnatural one; indeed it is one of the first to which we should think of applying Theorem I; compare Theorem II and its applications to which a reference is given.

smallest region of the kind already described for which z is a point of the locus Z. There is at least one of the points α_i on C. We shall show that all those points can be made to coincide on C.

Two points α_1 and α_2 which are on C can be made to coincide either on C or within C. For we fix z and the remainder of the points $\{\alpha_i\}$ and hence have a relation of type (2) between α_1 and α_2. A double point of the transformation is either on the circle C or interior to the region C.

Let us combine in this manner as many of the points $\{\alpha_i\}$ on C as possible, so as to leave the smallest number of points $\{\alpha_i\}$ actually on C. This number of points we denote by n; it is greater than zero and cannot be reduced.

We shall prove by induction that all these n points on C can be made to coalesce on C. This fact is evidently true for two points, for no double point of the corresponding transformation can be interior to C. We assume explicitly that whenever there are given $n-1$ points of this sort on an arc A of C, they can be made to coincide on A without changing any other point α_i, or the relation (1); we shall prove this fact for n points. It involves no loss of generality to suppose $n-1$ of these points at an end point X of A and the nth point at the other end point Y.

We consider one of the points α_1 at X and the point α_n at Y as connected by (1), while z and the other $\{\alpha_i\}$ remain fixed. When α_1 moves on C, α_n moves on C in the opposite sense, and the transformation (α_1, α_n) has a double point D interior to A. Consider α_1 and α_n to coincide at D. The new arc bounded by X and D contains the n points. We can make coincide, at a point X', $n-1$ points α_i, X' on the arc XD, and the other point α_n will be at D. Then we have the n points on an arc A' bounded by X' and Y', the arc A' shorter than A and every point of A' a point of A. Moreover, there are $n-1$ points at X', the end point nearest X, and the other is at $Y'=D$, the end point nearest Y.

We can continue in this manner successively to shorten our arc A, and we can shorten it indefinitely. For let us suppose we have a sequence of arcs (using the natural extension of our former notation), XY, $X'Y'$, $X''Y''$, $\cdots$, and suppose the points X, X', X'', $\cdots$ have a limit point x and the points Y, Y', Y'', $\cdots$ a limit point y different from x and of such a nature that we cannot make our arc of type A shorter than xy. We have $n-1$ coincident points α_i, successively approaching x, and one point α_n simultaneously approaching y. From the continuity of the left-hand member of (1) we can therefore suppose these $n-1$ points α_i to coincide at x and the other point α_n to lie at y. Then the procedure formerly used shows that we can bring the n points into an arc shorter than xy all of whose points are points of xy. This contradiction completes the proof of our statement that the arcs A, A',

A'', $\cdots$ shorten indefinitely, and hence the points $\alpha_1, \alpha_2, \cdots, \alpha_n$ can all be made to coincide at the limit point P of those arcs, with (1) still satisfied.

If $k = n$, Theorem I is proved. If $k > n$, consider the transformation (α_1, α_k) defined as before. We have supposed that none of the points $\alpha_1, \alpha_2, \cdots, \alpha_n$ can be moved from the circle C, so the transformation cannot contain α_k and hence α_k can be moved as near P as desired. Thus any other of the points $\alpha_{n+1}, \alpha_{n+2}, \cdots, \alpha_k$ can be moved as near P as desired without changing the value of $f(z)$. Hence all these points can be made to approach P and so the value of $f(z)$ is unchanged if all the $\alpha_1, \alpha_2, \cdots, \alpha_k$ coincide at P. Theorem I is now completely proved.

II. Special cases of Theorem I, with applications

There are two results, essentially special cases of Theorem I, which are particularly interesting in their applications.

Theorem II. *If the points $\alpha_1, \alpha_2, \cdots, \alpha_k$ lie in a circular region C and if z is exterior to C, the root of the equation in α*

$$\frac{1}{z-\alpha_1}+\frac{1}{z-\alpha_2}+\cdots+\frac{1}{z-\alpha_k}=\frac{k}{z-\alpha}$$

*lies in C.**

Theorem II is in reality a special case of the extension of Theorem I where we consider the quotient of two polynomials of the kind described. The denominator polynomial cannot vanish, since z is exterior to C.

For applications of Theorem II we refer to the citation already made and that made in the next following footnote. We proceed to another special case of Theorem I:

Theorem III. *If the points $\alpha_1, \alpha_2, \cdots, \alpha_k$ lie in a circular region C, the equation in α*

$$(z-\alpha_1)(z-\alpha_2)\cdots(z-\alpha_k)=(z-\alpha)^k \tag{4}$$

has at least one root in C.

Theorem III can be proved independently of Theorem I in a manner precisely analogous to the proof of Theorem II (loc. cit.). This proof of Theorem III involves a transcendental transformation of the α- (or (x, y)-) plane:

$$\alpha - z = x + iy = e^{u+iv},$$

and a study of the transform of C and certain centers of gravity in the (u, v)-plane. This proof gives in certain cases more detailed information than does Theorem I regarding the root α of (4). Thus if z is exterior to C, which is

* Walsh, these Transactions, vol. 22 (1921), p. 102; Lemma I.

Theorem II is closely connected with another more simple corollary of Theorem I, namely, that if k equal particles lie in a circle their center of gravity also lies in that circle.

the interior of a circle, we can write

$$k \log (z - \alpha) = \log (z - \alpha_1) + \log (z - \alpha_2) + \cdots + \log (z - \alpha_k),$$

where all the logarithms involved have the coefficients of $\sqrt{-1}$ in their pure imaginary parts lying between two numbers whose difference is less than π. We add the remark, without proof at this time, that there is no other finite region bounded by a regular curve which possesses the property of the finite circular region indicated by Theorem III.*

Theorem III can easily be extended to give some information concerning the location of *all* the roots of equation (4) in α. In fact, if α' is any root, the other roots are given by

$$(z - \alpha)^k = (z - \alpha')^k,$$
$$z - \alpha = \omega (z - \alpha'),$$

where ω is a kth root of unity. If we suppose that α' is in C, all the roots of (4) must lie in the k circular regions obtained by revolving C about z as center of rotation through the angles $2j\pi/k$, $j = 1, 2, \cdots, k$. In particular suppose one of these k circular regions, say C_1, is external to all the other $k - 1$ circular regions. Then we can prove that C_1 has on or within it precisely one root of (4). For consider the points $\alpha_1, \alpha_2, \cdots, \alpha_k$ to move continuously always remaining in C and to coincide at the center of C. In this situation C_1 contains precisely one root of (4). The roots of (4) vary continuously with the points $\alpha_1, \alpha_2, \cdots, \alpha_k$; none can enter or leave C_1 during motion of the kind indicated; so in the original situation C_1 contained precisely one root of (4).

However, our purpose is not primarily to study the roots of (4), but rather to use Theorems I and III as tools in proving more interesting relations. In preparation for these results we now prove the

LEMMA. *If the interiors and boundaries of the two circles C_1 and C_2, whose centers are α_1 and α_2 and radii r_1 and r_2 respectively, are the loci of two points z_1 and z_2, then as z_1 and z_2 vary independently, the locus of the point z which divides the segment (z_1, z_2) in the constant ratio $(m_1 : m_2)$,*

$$z = \frac{m_2 z_1 + m_1 z_2}{m_1 + m_2} \qquad (m_1 \neq -m_2, \quad m_1 m_2 \neq 0), \tag{5}$$

is the interior and boundary of a circle C whose center is

$$\frac{m_2 \alpha_1 + m_1 \alpha_2}{m_1 + m_2}$$

and whose radius is

$$\left| \frac{m_2 r_1}{m_1 + m_2} \right| + \left| \frac{m_1 r_2}{m_1 + m_2} \right|.$$†

* For the corresponding fact for Theorem II, compare Walsh, these Transactions, vol. 24 (1922), pp. 31–69; Theorem III.

† The terminology that z divides the segment (z_1, z_2) in the ratio $(m_1 : m_2)$ is usual when

If z_1 and z_2 are on or within C_1 and C_2 respectively,

$$|z_1 - \alpha_1| \leqq r_1, \qquad |z_2 - \alpha_2| \leqq r_2,$$

we have

$$z - \frac{m_2\,\alpha_1 + m_1\,\alpha_2}{m_1 + m_2} = \frac{m_2}{m_1 + m_2}(z_1 - \alpha_1) + \frac{m_1}{m_1 + m_2}(z_2 - \alpha_2),$$

which is in absolute value less than or equal to

$$\left|\frac{m_2\,r_1}{m_1 + m_2}\right| + \left|\frac{m_1\,r_2}{m_1 + m_2}\right|.$$

It remains to be shown that given any point z on or within C, we can properly determine z_1 and z_2. In order to do this, we merely place

$$\frac{m_2(z_1 - \alpha_1)}{m_1 + m_2} = \left(z - \frac{m_2\,\alpha_1 + m_1\,\alpha_2}{m_1 + m_2}\right)\frac{\left|\dfrac{m_2\,r_1}{m_1 + m_2}\right|}{\left|\dfrac{m_2\,r_1}{m_1 + m_2}\right| + \left|\dfrac{m_1\,r_2}{m_1 + m_2}\right|},$$

$$\frac{m_1(z_2 - \alpha_2)}{m_1 + m_2} = \left(z - \frac{m_2\,\alpha_1 + m_1\,\alpha_2}{m_1 + m_2}\right)\frac{\left|\dfrac{m_1\,r_2}{m_1 + m_2}\right|}{\left|\dfrac{m_2\,r_1}{m_1 + m_2}\right| + \left|\dfrac{m_1\,r_2}{m_1 + m_2}\right|}.$$

Then z_1 and z_2 lie in the proper regions and satisfy (5), so the proof of the lemma is complete.

This lemma gives no result if $m_1 + m_2 = 0$; to treat this case we take (5) in the form

$$m_2(z - z_1) = -\,m_1(z - z_2), \tag{6}$$

which is equivalent for our present purpose. If C_1 and C_2 are mutually external there is no point z different from the point at infinity which satisfies (6). If C_1 and C_2 are not mutually external we may choose $z_1 = z_2$, and every point of the plane satisfies (6). In the theorems below we give the general formulas for the determination of C with the understanding that when $m_1 + m_2 = 0$, C is considered to contain no finite point of the plane or every finite point of the plane according as C_1 and C_2 are or are not mutually external.*

m_1 and m_2 are real; whether m_1 and m_2 are real or not, we simply understand that statement to mean that z is given by (5).

This lemma was proved by the present writer and by the present method for m_1 and m_2 real and positive, in a note in the *Comptes Rendus du Congrès international*, Strasbourg, 1920. See also Theorem II of the reference given in connection with the present Theorem II.

* Thus, a careful statement of Theorem IV for the case $A = 1$ is

THEOREM IVa. *If the points $a_1, a_2, \cdots, a_k$ have as their locus a circle C_1 whose center is α and radius r_1, and if the points $b_1, b_2, \cdots, b_k$ have as their locus a circle C_2 whose center is β*

We now apply Theorem III and the Lemma to prove

THEOREM IV. *If the points $a_1, a_2, \cdots, a_k$ have as their locus (the interior and boundary of) a circle C_1 whose center is α and radius r_1, and if the points $b_1, b_2, \cdots, b_k$ have as their locus a circle C_2 whose center is β and radius r_2, the roots of the equation*

$$(7) \quad (z-a_1)(z-a_2)\cdots(z-a_k) - A(z-b_1)(z-b_2)\cdots(z-b_k) = 0$$

have as their locus the k circles with the respective radii

$$\frac{r_1 + |A^{1/k}| r_2}{|1 - A^{1/k}|}$$

and centers

$$(8) \qquad \frac{\alpha - A^{1/k}\beta}{1 - A^{1/k}}$$

where $A^{1/k}$ takes all the k values possible. Any one of these k circles which is external to all the others contains precisely one root of (7).

It is evident that any polynomial equation can be written in the form (7), and in an infinite variety of ways.

We prove Theorem IV as follows. If a point z is a root of (7), an application of Theorem III shows that

$$(z-a)^k - A(z-b)^k = 0,$$

where a and b are some points on or within C_1 and C_2 respectively. Hence we have

$$z - a = A^{1/k}(z-b),$$

$$z = \frac{a - A^{1/k} b}{1 - A^{1/k}},$$

where $A^{1/k}$ is some kth root of A, so z is on or within one of the k circles of the theorem. Variation of a and b independently and over C_1 and C_2 as loci shows that the k circles are the actual loci of the roots of (7). Continuous motion of the points $a_1, a_2, \cdots, a_k, b_1, b_2, \cdots, b_k$ so as to remain in their proper regions

and radius r_2, and if C_1 and C_2 are mutually external, then the roots of the equation

$$(7') \qquad (z-a_1)(z-a_2)\cdots(z-a_k) = (z-b_1)(z-b_2)\cdots(z-b_k)$$

have as their locus the $k-1$ circles with the respective radii

$$\frac{r_1 + r_2}{|1 - \omega|}$$

and centers

$$\frac{\alpha - \omega\beta}{1 - \omega}$$

where ω takes the $k-1$ values of the kth roots of unity which differ from unity itself. Any one of these $k-1$ circles which is external to all the others contains precisely one root of (7').

If C_1 and C_2 are not mutually external, the locus of the roots of (7') is the entire plane.

The details of the proof of Theorem IV*a* are similar to those of the proof of Theorem IV and are left to the reader.

and finally to coincide shows in precisely the manner used in considering the roots of (4) that any one of the k circles which is exterior to all the others contains precisely one root of (7).

Theorem IV reduces to the lemma when $k = 1$.

A few remarks on the geometry of the situation are not out of place. The k points (8) all lie on the circle which is the locus of points z such that

$$\left|\frac{z-\alpha}{z-\beta}\right| = |A^{1/k}|,$$

which circle is of course a circle of the coaxial family determined by α and β as null circles. Each of the k points also lies on a circle which is the locus of points z such that

$$\frac{\dfrac{z-\alpha}{|z-\alpha|}}{\dfrac{z-\beta}{|z-\beta|}} = \frac{A^{1/k}}{|A^{1/k}|};$$

that is, the argument of the ratio $(z-\alpha)/(z-\beta)$ is constant. These k circles belong to the coaxial family of all circles passing through α and β, the family conjugate to the former coaxial family mentioned. These same k circles can be arranged in order so that at α and β each cuts its predecessor at an angle of $2\pi/k$.

It is quite easy for us to obtain results concerning the roots of the derivatives of equation (7). The mth derivative $(k > m \geqq 0)$, except for a constant factor, can be written

$$(9)\quad (z-a_1')(z-a_2')\cdots(z-a_{k-m}') - A(z-b_1')(z-b_2')\cdots(z-b_{k-m}') = 0.$$

The points a_j' all lie on or within C_1 and the points b_j' all lie on or within C_2, by the theorem of Lucas concerning the roots of the derivative of a polynomial. Hence all the roots of (9) lie on or within the $k-m$ circles with the respective radii

$$\frac{r_1 + |A^{1/(k-m)}|\, r_2}{|1 - A^{1/(k-m)}|}$$

and centers

$$\frac{\alpha - A^{1/(k-m)}\beta}{1 - A^{1/(k-m)}},$$

where $A^{1/(k-m)}$ takes all the $k-m$ values possible. These circles form the locus of the roots of (9). Any one of these $k-m$ circles which is external to the remaining $k-m-1$ circles contains precisely one root of (9). This new result includes Theorem IV as the case $m = 0$.

We now state another theorem which results from Theorem III and which like Theorem IV refers to the sum of two polynomials.*

THEOREM V. *If the points $a_1, a_2, \cdots, a_k$ lie on or within a circle C whose center is α and radius r, all the roots of the equation*

$$(z - a_1)(z - a_2) \cdots (z - a_k) - A = 0 \tag{10}$$

lie on or within one of the k circles which have the common radius r and whose centers are the points $\alpha + A^{1/k}$, where $A^{1/k}$ takes all the k possible values. If these k circles are mutually external, each contains precisely one root of (10).

Of course any polynomial equation can be written in the form (10) and in an infinite variety of ways. The proof of the theorem follows the proof of Theorem IV. If C is the locus of $a_1, a_2, \cdots, a_k$, the k circles form the locus of the roots of (10).

Equation (10) is particularly interesting because the derived equation is independent of A. All the roots of the mth derived equation, $m < k$, lie on or within C. Theorem V can also be extended to a polynomial of the form

$$\frac{d}{dz}[(z - a_1)(z - a_2) \cdots (z - a_k)] - A = 0,$$

but this generalization as well as Theorem V itself is included in a more general theorem to be proved from Theorem I.

III. SOME DIRECT APPLICATIONS OF THEOREM I

We now proceed to derive a number of results directly from Theorem I instead of from Theorem II or Theorem III.

If in Theorem I the polynomial $f(z)$ contains merely one set of variables $\{\alpha_i\}$ referring to a single circular region, the locus Z can have as its boundary no point other than a point of the curve traced by z when $\alpha = \alpha_1 = \alpha_2 = \cdots = \alpha_k$ traces the circle C_α. This leads to

THEOREM VI. *Let the common locus of all the roots of a polynomial*

$$f(z) = (z - \alpha_1)(z - \alpha_2) \cdots (z - \alpha_k)$$

be the interior and boundary of a circle C whose center is α and radius r. Then the locus of the roots of the polynomial

$$A_0 f(z) + A_1 f'(z) + \cdots + A_{k-1} f^{(k-1)}(z) + A_k f^{(k)}(z) \tag{11}$$

is composed of the interiors and boundaries of the circles whose common radius is r and whose centers are the roots of

$$A_0 (z - \alpha)^k + kA_1 (z - \alpha)^{k-1} + \cdots + k(k-1) \cdots 2 \cdot 1 A_k. \tag{12}$$

* If in Theorem IV we replace (7) by the sum of two polynomials not of the same degree and neither of degree zero, we are led to the determination of a locus which is not generally bounded by circles. Compare 4 of the next to the last paragraph of this paper.

Any of these circles having no point in common with any other of these circles contains a number of roots of (11) *equal to the multiplicity of the center of that circle as a root of* (12).

A detailed proof of this theorem is quite simple. For any particular root z of (11) we consider the points $\{\alpha_i\}$ to coincide at $\overline{\alpha}$ on or in C. We have then

$$A_0(z-\overline{\alpha})^k + kA_1(z-\overline{\alpha})^{k-1} + \cdots + k(k-1)\cdots 2\cdot 1A_k = 0.$$

Since $|\overline{\alpha}-\alpha| \leqq r$, we know that z must lie in or on one of the circles determined. Conversely, any point z in or on one of these circles is a root of (11) for some choice of the $\{\alpha_i\}$; denote the center of the circle by β. We need merely choose $\alpha_1 = \alpha_2 = \cdots = \alpha_k = \alpha + z - \beta$, which is a point on or within C.

If one of the circles C_1 has no point in common with any other of those circles, we may vary the α_i at will and no root of (11) can enter or leave C_1; these roots vary continuously if the α_i vary continuously. In particular if

$$\alpha_1 = \alpha_2 = \cdots = \alpha_k = \alpha,$$

we have the number of roots of (11) proper to C_1. It is similarly true that a number of these circles $C_1, \cdots, C_l$ having no point in common with any other of the circles contain a number of roots of (11) equal to the sum of the multiplicities of their centers considered as roots of (12).

Theorem VI enables us to give a result concerning the location of the roots of the sum of any number of arbitrary polynomials; $f(z)$ is to be chosen one of those polynomials of highest degree.

If in Theorem VI we place $A_0 = A_2 = A_3 = \cdots = A_k = 0$, $A_1 \neq 0$, we have

A circle which contains all the roots of a polynomial contains also all the roots of its derivative.

This is essentially the theorem of Lucas:

*A convex polygon which contains all the roots of a polynomial contains also all the roots of its derivative.**

The similar results obtained from Theorem VI by placing $A_l \neq 0$ but all the other A_i zero $(l > 0)$ give the theorem of Lucas for the other derivatives of $f(z)$.

Theorem I applies just as well to the integral of a polynomial as to its derivative, but for a result of the nature of those just given it is necessary to choose a *particular* integral. Thus we shall prove

THEOREM VII. *Let the common locus of all the roots of the polynomial*

$$f(z) = (z-\alpha_1)(z-\alpha_2)\cdots(z-\alpha_k)$$

* In both of these statements as well as many other similar results, the term *contains* may be interpreted to include the possibility of points on the boundary or to exclude that possibility. Whichever interpretation is considered in the hypothesis, the conclusion will be true under that same interpretation.

be the interior and boundary of the circle C whose center is α and radius r. Then the locus of the roots of the polynomial

$$F(z) = \int_0^z f(z)\,dz$$

is composed of the interiors and boundaries of the $k+1$ circles whose centers are $(1-\omega)\alpha$ and radii $|1-\omega|r$, where ω takes all the values of the $(k+1)$st roots of unity. Any of these circles which is entirely exterior to all the others contains precisely one root of $F(z)$.

If z is a root of $F(z)$ for particular values of the α_i, we must have for some choice of $\bar{\alpha} = \alpha_1 = \alpha_2 = \cdots = \alpha_k$ in or on C,

$$F(z) = \frac{(z-\bar{\alpha})^{k+1}}{k+1} - \frac{(-\bar{\alpha})^{k+1}}{k+1} = 0,$$
$$z = (1-\omega)\bar{\alpha},$$

where $\omega^{k+1} = 1$. From this equation we have at once the desired inequality:

$$z - (1-\omega)\alpha = (1-\omega)(\bar{\alpha} - \alpha),$$
$$|z - (1-\omega)\alpha| \leqq |1-\omega|r,$$

which shows that z is on or within one of the $k+1$ circles. On the other hand, any point z on or within one of these $k+1$ circles is a root of $F(z)$ for proper choice of the $\{\alpha_i\}$. In fact, if ω is the particular root of unity corresponding to that circle we have merely to place $\bar{\alpha} = \alpha_1 = \alpha_2 = \cdots = \alpha_k$, where

$$\bar{\alpha} - \alpha = \frac{z - (1-\omega)\alpha}{1-\omega};$$

the exceptional case $\omega = 1$ is trivial.

The proof of the theorem is now complete except for the last sentence; this is proved in precisely the same manner as the corresponding statement in Theorem VI. Like Theorem VI, Theorem VII illustrates the remark made just previous to the statement of Theorem VI.

We shall next prove further results concerning the roots of the derivative of a polynomial, making continued use of the Lemma. The special case of the Lemma where m_1 and m_2 are positive is particularly simple; this special case leads to

THEOREM VIII. *Denote by $g(z)$ the polynomial*

$$(z-z_1)^{m_1}(z-z_2)^{m_2}$$

and by $m_1^{(n)} : m_2^{(n)}$ $(n = 1, 2, \cdots, m)$ the ratios in which the m distinct roots $z^{(n)}$ of $g^{(k)}(z)$ (the kth derivative of $g(z)$) divide the segment (z_1, z_2). Let the interiors and boundaries of circles C_1 and C_2 whose centers are α_1 and α_2 and radii r_1 and r_2 be the loci respectively of m_1 and m_2 roots of a polynomial $f(z)$ which*

* The reader will easily prove from Rolle's Theorem that no point $z^{(n)}$ distinct from z_1 and z_2 can be a multiple root of $g^{(k)}(z)$.

has no other roots. Then the locus of the roots of $f^{(k)}(z)$, the kth derivative of $f(z)$, is composed of the m circles $C^{(n)}$ whose centers are

$$\frac{m_2^{(n)}\alpha_1 + m_1^{(n)}\alpha_2}{m_1^{(n)} + m_2^{(n)}}$$

and whose radii are

$$\frac{m_2^{(n)} r_1 + m_1^{(n)} r_2}{m_1^{(n)} + m_2^{(n)}}.$$

If one of these circles $C^{(n)}$ is exterior to all the others, it contains a number of roots of $f(z)$ equal to the multiplicity of $z^{(n)}$ as a root of $g^{(k)}(z)$.

The proof of this theorem is quite similar to the proofs of Theorems VI and VII and will therefore be omitted.*

If a number of circles $C^{(n)}$ of Theorem VIII overlap and are entirely exterior to all the other circles $C^{(n)}$, they contain together a number of roots of $f^{(k)}(z)$ equal to the sum of the multiplicities of the corresponding roots $z^{(n)}$ of $g^{(k)}(z)$. A similar situation for one as for several circles arises in connection with Theorems VI and VII. In the future we shall indicate this general fact by saying that the number of roots of $f^{(k)}(z)$ *proper* to a circle $C^{(n)}$ is the multiplicity of the point $z^{(n)}$ as a root of $g^{(k)}(z)$.

In Theorem VIII we are in reality considering a polynomial

$$f(z) = \phi(z) \cdot \psi(z),$$

where the m_1 roots of $\phi(z)$ lie on or within C_1 and the m_2 roots of $\psi(z)$ lie on or within C_2. Then our conclusion refers to the polynomial

$$f^{(k)}(z) = \phi \cdot \psi^{(k)} + \frac{k}{1!}\phi' \cdot \psi^{(k-1)} + \frac{k(k-1)}{2!}\phi'' \cdot \psi^{(k-2)} + \cdots + \phi^{(k)} \cdot \psi.$$

When the result is expressed in this form it can be given a large extension. Under the same hypothesis with respect to ϕ and ψ we consider the polynomial

$$(13) \quad A_0\phi \cdot \psi^{(k)} + A_1\phi' \cdot \psi^{(k-1)} + A_2\phi'' \cdot \psi^{(k-2)} + \cdots + A_k\phi^{(k)} \cdot \psi.$$

If z is a root of (13) for a particular choice of the roots of ϕ and ψ, there exist α and β in or on C_1 and C_2 respectively and such that

$$\begin{aligned}
&m_2(m_2-1)\cdots(m_2-k+1)A_0(z-\alpha)^{m_1}(z-\beta)^{m_2-k} \\
&\quad + m_1 \cdot m_2(m_2-1)\cdots(m_2-k+2)A_1(z-\alpha)^{m_1-1}(z-\beta)^{m_2-k+1} \\
&\quad + m_1 \cdot (m_1-1) \cdot m_2 \cdot (m_2-1) \\
&\qquad \cdots (m_2-k+3)A_2(z-\alpha)^{m_1-2}(z-\beta)^{m_2-k+2} \\
&\quad \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
&\quad + m_1(m_1-1)\cdots(m_1-k+1)A_k(z-\alpha)^{m_1-k}(z-\beta)^{m_2} = 0.
\end{aligned} \tag{14}$$

* This theorem was published in a short note by the present writer, Paris Comptes Rendus, vol. 172 (1921), pp. 662–664, and its proof as there indicated contains the germ of the proof of Theorem I. The special case of Theorem VIII for the case $k = 1$ had been previously proved by means of Theorem II; see the reference given in connection with that theorem.

Equation (14) is essentially an equation in homogeneous coördinates μ_1, μ_2:

$$
\begin{aligned}
&m_2(m_2-1)\cdots(m_2-k+1)A_0\,\mu_1^{m_1}\mu_2^{m_2-k}\\
&\quad+m_1\cdot m_2\cdot(m_2-1)\cdots(m_2-k+2)A_1\,\mu_1^{m_1-1}\mu_2^{m_2-k+1}\\
&\quad+m_1(m_1-1)\cdot m_2\cdot(m_2-1)\cdots(m_2-k+3)A_2\,\mu_1^{m_1-2}\mu_2^{m_2-k+2}\\
&\qquad\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\
&\quad+m_1(m_1-1)\cdots(m_1-k+1)A_k\,\mu_1^{m_1-k}\mu_2^{m_2}=0,
\end{aligned}
\tag{15}
$$

whose distinct solutions, m in number, we denote by $\mu_1^{(n)}:\mu_2^{(n)}$, $n=1,2,\cdots,m$. From (14) and (15) we see that for some n we have

$$
\frac{z-\alpha}{z-\beta}=\frac{\mu_1^{(n)}}{\mu_2^{(n)}},\qquad z=\frac{\mu_2^{(n)}\alpha+\mu_1^{(n)}\beta}{\mu_1^{(n)}+\mu_2^{(n)}}.
$$

The Lemma immediately gives us certain circles in at least one of which must be located the point z. It is easily seen that these circles give the exact locus of z, and we have

THEOREM IX. *Let the interiors and boundaries of circles C_1 and C_2, whose centers are α_1 and α_2 and radii r_1 and r_2, be respectively the loci of the m_1 roots of a polynomial ϕ and the m_2 roots of a polynomial ψ. Then the locus of the roots of* (13) *is composed of the interiors and boundaries of the m circles $C^{(n)}$ whose centers are the points*

$$
\frac{\mu_2^{(n)}\alpha_1+\mu_1^{(n)}\alpha_2}{\mu_1^{(n)}+\mu_2^{(n)}}
$$

and radii

$$
\frac{|\mu_2^{(n)}|r_1+|\mu_1^{(n)}|r_2}{|\mu_1^{(n)}+\mu_2^{(n)}|}\qquad(n=1,2,\cdots,m),
$$

where $\mu_1^{(n)}:\mu_2^{(n)}$ are the m distinct roots of (15). *The number of roots of* (13) *proper to a circle $C^{(n)}$ is the multiplicity of the ratio $\mu_1^{(n)}:\mu_2^{(n)}$ as a root of* (15).

The methods that we have been using enable us also to obtain a result for the roots of the derivative of an entire transcendental function of the simple type

$$
f(z)=e^{P(z)}\cdot Q(z)
$$

where

$$
\begin{aligned}
P(z)&=a(z-\alpha_1)(z-\alpha_2)\cdots(z-\alpha_p),\qquad a\neq 0,\\
Q(z)&=b(z-\beta_1)(z-\beta_2)\cdots(z-\beta_q),\qquad b\neq 0.
\end{aligned}
$$

The roots of $f'(z)$ are given by the equation

$$
Q(z)P'(z)+Q'(z)=0.\tag{16}
$$

If all the α_i and β_i lie in or on a circle C whose center is α and radius r, and if z is a root of (16), we shall have for some $\overline{\alpha}$ and $\overline{\beta}$ in or on C

$$
ap(z-\overline{\beta})^q(z-\overline{\alpha})^{p-1}+q(z-\overline{\beta})^{q-1}=0.
$$

Then we must have either

$$(z - \overline{\beta})^{q-1} = 0$$

or

$$ap(z - \overline{\beta})(z - \overline{\alpha})^{p-1} + q = 0.$$

If this latter equation is verified we find from Theorem III that there exists a point $\overline{\gamma}$ in or on C such that

$$ap(z - \overline{\gamma})^p + q = 0,$$

$$z - \overline{\gamma} = \omega \sqrt[p]{\frac{-q}{ap}}, \qquad \text{where} \qquad \omega^p = 1.$$

It therefore follows that

$$z - \left(\alpha + \omega \sqrt[p]{\frac{-q}{ap}}\right) = \overline{\gamma} - \alpha,$$

which is less than or equal to r in absolute value. Then all the roots of $f'(z)$ lie in C and the p circles of common radius r whose centers are the points

$$\alpha + \omega \sqrt[p]{\frac{-q}{ap}}.$$

We state this fact and its obvious converse in

THEOREM X. *Let the interior and boundary of the circle C whose center is α and radius r be the common locus of all the roots of the two polynomials*

$$P(z) = a(z - \alpha_1)(z - \alpha_2) \cdots (z - \alpha_p), \qquad a \neq 0,$$

$$Q(z) = b(z - \beta_1)(z - \beta_2) \cdots (z - \beta_q), \qquad b \neq 0.$$

Then the locus of the roots of the derivative of the function

$$e^{P(z)} \cdot Q(z)$$

is composed of C (unless $q = 1$) and the p circles whose common radius is r and whose centers are the points

$$\alpha + \omega \sqrt[p]{\frac{-q}{ap}},$$

where ω takes the values of the pth roots of unity. There are $q - 1$ roots of the derivative proper to C and one proper to each of the other p circles.

In the proof of this theorem we have used not the fact that the interior and boundary of C is the locus of the roots of $P(z)$ but that it is the locus of the roots of $P'(z)$, which gives in reality a more general result. Thus if $p = 1$, there is no restriction whatever on the root of $P(z)$.

We now prove a theorem which is a generalization of our result concerning the roots of (16) corresponding to our former generalization of Theorem VIII.

THEOREM XI. *Let the interior and boundary of a circle C whose center is α and radius is r be the common locus of the roots of the polynomials $\phi_1, \phi_2, \cdots, \phi_q$ of respective degrees $k_1, k_2, \cdots, k_q$. Then the locus of the roots of the polynomial*

$$(17) \qquad \phi_1 \cdot \phi_2 \cdot \phi_3 \cdot \cdots \cdot \phi_q - A\phi_1^{(n_1)} \cdot \phi_2^{(n_2)} \cdot \cdots \cdot \phi_q^{(n_q)}, \qquad n_i \leqq k_i,$$

is composed of the circle C (unless $n_i = k_i$ for $i = 1, 2, \cdots, q$) and the circles, $n_1 + n_2 + \cdots + n_q$ in number, whose centers are the points

$$\alpha + \omega[k_1 \cdot (k_1 - 1) \cdots (k_1 - n_1 + 1) \cdot k_2 \cdot (k_2 - 1) \cdots (k_2 - n_2 + 1) \cdots k_q \cdot (k_q - 1) \cdots (k_q - n_q + 1) \cdot A]^{1/(n_1+n_2+\cdots+n_q)}$$

where ω takes all possible values of an $(n_1 + n_2 + \cdots + n_q)$th root of unity, and whose common radius is r. There are

$$(k_1 + k_2 + \cdots + k_q) - (n_1 + n_2 + \cdots + n_q)$$

roots of (17) *proper to C and one root proper to each of the $n_1 + n_2 + \cdots + n_q$ other circles.*

Under the hypothesis of the theorem, if z is a root of (17) there exist points $\alpha', \alpha'', \cdots, \alpha^{(q)}$ on or within C and such that we have

$$(z - \alpha')^{k_1}(z - \alpha'')^{k_2} \cdots (z - \alpha^{(q)})^{k_q} - [k_1 \cdot (k_1 - 1) \cdots (k_1 - n_1 + 1) \cdot k_2 \cdot (k_2 - 1) \cdots (k_2 - n_2 + 1) \cdot k_q \cdot (k_q - 1) \cdots (k_q - n_q + 1) \cdot A \cdot (z - \alpha')^{k_1 - n_1}(z - \alpha'')^{k_2 - n_2} \cdots (z - \alpha^{(q)})^{k_q - n_q}] = 0.$$

Then z must satisfy one of the equations

$$(z - \alpha')^{k_1 - n_1}(z - \alpha'')^{k_2 - n_2} \cdots (z - \alpha^{(q)})^{k_q - n_q} = 0,$$

$$(z - \alpha')^{n_1}(z - \alpha'')^{n_2} \cdots (z - \alpha^{(q)})^{n_q} - k_1(k_1 - 1) \cdots (k_q - n_q + 1)A = 0.$$

If z satisfies this latter equation we know from Theorem III that there exists a point $\bar{\alpha}$ on or within C and such that

$$(z - \bar{\alpha})^{n_1+n_2+\cdots+n_q} = k_1(k_1 - 1) \cdots (k_q - n_q + 1)A.$$

That is, we have the equation

$$z - \{\alpha + \omega[k_1(k_1 - 1) \cdots (k_q - n_q + 1)A]^{1/(n_1+n_2+\cdots+n_q)}\} = \bar{\alpha} - \alpha,$$

where $\omega^{n_1+n_2+\cdots+n_q} = 1$. The absolute value of the right-hand member is not greater than r, so z must lie in or on one of the $n_1 + n_2 + \cdots + n_q$ circles of the theorem.

The detailed proof of the remainder of Theorem XI now requires no further analytical work and is left to the reader.

The present paper aims merely to give some of the more immediate results that can be obtained from Theorem I. Further results can be obtained by extending in various directions:

1. Application of these methods to other and more general types of polynomials.

2. Extension of the results to other circular regions, namely, half-planes and the exteriors of circles.

3. Expression of the results in a manner which shall be independent of linear transformation of the complex variable.

4. Detailed determination and study of loci which naturally arise and which are of the same general nature as the locus determined in the lemma. For example, if the loci of α and β are two circular regions, determine the locus of the points z defined by

$$(z - \alpha)^m = A(z - \beta)^n.$$

For only the simplest polynomials $f(z, \alpha_i, \beta_i)$ do we get a number of circular regions as the locus of the roots of f when the loci of the roots of the α_i and β_i are circular regions. All the problems of the present paper have been characterized by a certain *linearity*, either the original problems themselves or those problems after a simple transformation. In more general problems, where the locus of the roots does not ordinarily consist of circular regions, the exact locus should be determined and also a simple even if rough approximation.

5. Application of these results and methods to the case of real polynomials.*

The present writer hopes to return later to a consideration of these questions.

* Thus the following theorem results from Theorem VI:

Let all the roots of a polynomial $f(z)$ of degree k lie in an interval of the axis of reals whose center is α and length l. Denote the roots of the polynomial

$$A_0 z^k + kA_1 z^{k-1} + \cdots + k(k-1)\cdots 2\cdot 1\cdot A_k \tag{1'}$$

by $z_1, z_2, \cdots, z_k$, which roots are supposed real. Then all the roots of the polynomial

$$A_0 f + A_1 f' + \cdots + A_k f^{(k)} \tag{2'}$$

lie in the intervals of the axis of reals whose centers are the points $\alpha + z_i$ $(i = 1, 2, \cdots, k)$ and whose common length is l. The number of roots of (2') proper to each of these intervals is the multiplicity of the corresponding z_i as a root of (1').

HARVARD UNIVERSITY,
CAMBRIDGE, MASS. (JUNE, 1921.)

A NOTE ON THE PRECEDING PAPER*

BY

D. R. CURTISS

The results of Dr. Walsh's paper are so interesting and capable of so many applications that it may be worth while to indicate a briefer proof of his fundamental theorem which connects with earlier work in this field.

First let us note that Walsh's Theorem II is a form of Laguerre's Theorem given below. In fact, if we write

$$f(z) = (z - \alpha_1)(z - \alpha_2) \cdots (z - \alpha_k),$$

Walsh's relation

$$\frac{1}{z - \alpha_1} + \frac{1}{z - \alpha_2} + \cdots + \frac{1}{z - \alpha_k} = \frac{k}{z - \alpha}$$

becomes

$$\frac{f'(z)}{f(z)} = \frac{k}{z - \alpha},$$

or

$$\alpha = z - k\frac{f(z)}{f'(z)}. \tag{1}$$

Walsh's Theorem II states that if the roots of $f(z)$ are in the circular region C, and if z is exterior to C, then α lies in C. This is precisely the result obtained by Laguerre as given on page 59 of volume I of his collected works, if expressed in non-homogeneous coördinates as on page 57.

We shall use another form of Laguerre's Theorem (loc. cit., p. 57) to prove the underlying proposition on which Walsh bases his proof of his Theorem I, as follows:

Every circle through any point z and its "derived point" α as defined by (1) *either passes through all the roots of $f(z)$, or else has at least one root in the region interior to it, and at least one root in the exterior region.*

I have recently called attention† to the fact that this is a corollary of a theorem of Bôcher's on jacobians which has served as a starting point in Walsh's earlier papers.

With a slight change in Walsh's notation, whereby we substitute z's for α's, the proposition from which Theorem I is deduced may be stated thus:

* Presented to the Society April 14, 1922, under the title *On the zeros of successive polars of a binary form.*

† Science, vol. 55 (1922), p. 193.

Let $z_1, z_2, \cdots, z_k$ lie within a circular region C, and let $\sum_r^k$ denote the sum of all products of the z's taken r at a time. Then the equation

$$(2) \qquad a_0 z^k + k a_1 z^{k-1} + \frac{k(k-1)}{2!} a_2 z^{k-2} + \cdots + k a_{k-1} z + a_k = 0,$$

where

$$-a_k = a_0 \sum_k^k + a_1 \sum_{k-1}^k + a_2 \sum_{k-2}^k + \cdots + a_{k-1} \sum_1^k$$

has at least one root in C for each set of values $a_0, a_1, \cdots, a_{k-1}$.

This theorem is trivial if all the a's are zero. In any other case we proceed by induction, assuming our proposition true for $k = r - 1$ and proving its validity for $k = r$. It is obviously true for $k = 1$.

Let us place $k = r$ in (2) and in (1). The resulting left-hand member of (2) we will use as $f(z)$ in (1). We write

$$f(z) = a_0 z^r + r a_1 z^{r-1} + \cdots + a_r,$$

$$z f'(z) - r f(z) = -r[a_1 z^{r-1} + (r-1) a_2 z^{r-2} + \cdots + a_r],$$

so that (1) becomes

$$(3) \qquad \alpha = -\frac{a_1 z^{r-1} + (r-1) a_2 z^{r-2} + \cdots + a_r}{a_0 z^{r-1} + (r-1) a_1 z^{r-2} + \cdots + a_{r-1}}.$$

Now consider the equation obtained by replacing α in (3) by z_r; cleared of fractions and arranged according to powers of z this is

$$(4) \quad (a_0 z_r + a_1) z^{r-1} + (r-1)(a_1 z_r + a_2) z^{r-2} + \cdots + (a_{r-1} z_r + a_r) = 0.$$

But if we compare this with (2), and note that

$$\begin{aligned} -(a_{r-1} z_r + a_r) &= -a_{r-1} z_r + a_0 \sum_r^r + a_1 \sum_{r-1}^r \\ &\qquad + a_2 \sum_{r-2}^r + \cdots + a_{r-1} \sum_1^r \\ &= -a_{r-1} z_r + a_0 z_r \sum_{r-1}^{r-1} + a_1 (z_r \sum_{r-2}^{r-1} + \sum_{r-1}^{r-1}) \\ &\qquad + a_2 (z_r \sum_{r-3}^{r-1} + \sum_{r-2}^{r-1}) + \cdots + a_{r-1} (z_r + \sum_1^{r-1}) \\ &= (a_0 z_r + a_1) \sum_{r-1}^{r-1} + (a_1 z_r + a_2) \sum_{r-2}^{r-1} \\ &\qquad + \cdots + (a_{r-2} z_r + a_{r-1}) \sum_1^{r-1}, \end{aligned}$$

we see that (4) is a case of (2) for $k = r - 1$, and hence (4) has a root Z_r in C.

We now give z the value Z_r in (3) and obtain the result that the "derived point" α is z_r. By Laguerre's Theorem every circle through Z_r and z_r either has on it a root of $f(z)$ or has a root in its interior and one exterior to it; but Z_r and z_r are in C, so that we can draw through them a circle interior to C. Thus $f(z)$ has a root in C, and on completing our induction we see that (2) has a root in C.

The theorem we have just proved can not, however, be regarded as a new one. It is equivalent to the theorem of Grace:*

If a form $f(z)$ is apolar to a given form $\phi(z)$, then it has a zero within any circle enclosing all the roots of $\phi(z)$.

If equation (2) and the equation

$$(z - z_1)(z - z_2) \cdots (z - z_k) = 0 \tag{5}$$

are made homogeneous, the condition that (2) and (5) be apolar is precisely the relation just following (2). Since this holds, Grace's theorem requires that a root of (2) lie within any circle enclosing all the roots of (5), which are $z_1, z_2, \cdots, z_k$.

Grace's theorem is a consequence of Laguerre's when the latter is put in homogeneous form. If in (1) we replace α by ξ/η, z by x/y, and $f(z)$ by $y^k f(x, y)$, (1) takes the form

$$\xi f'_x + \eta f'_y = 0, \tag{6}$$

which is the equation of the first polar of (ξ, η) with respect to $f(x, y)$. One form of Laguerre's theorem (loc. cit., p. 59) may, then, be thus stated:

If the point $\alpha = \xi/\eta$ is without a circle C containing all the roots of f, all the roots of the first polar of α with respect to f are within C.

Consider a set of points $z_1, z_2, \cdots, z_k$, and let f_1 be the first polar of z_1 with respect to f, let f_2 be the first polar of z_2 with respect to f_1, etc. If all these z's are outside C, the roots of all the polynomials $f_1, f_2, \cdots, f_{k-1}$ are within C. If (2) and (5) are apolar, z_k is the root of f_{k-1}, and Grace's theorem, in which f and ϕ may be interchanged, follows directly.

The "mixed polars" $f_1, f_2, \cdots$ are perhaps worth writing down in another form. Thus if we use the notation, for $r < k$,

$$\psi(z) = (z - z_1)(z - z_2) \cdots (z - z_r) = c_0 z^r + c_1 z^{r-1} + \cdots + c_r,$$

we have

$$(-1)^r f_r(x, y) = c_r \frac{\partial^r f}{\partial x^r} - c_{r-1} \frac{\partial^r f}{\partial x^{r-1} \partial y} + \cdots + (-1)^r c_0 \frac{\partial^r f}{\partial y^r}.$$

In non-homogeneous notation the equation $f_r = 0$ becomes

* *The zeros of a polynomial*, Proceedings of the Cambridge Philosophical Society, vol. 11 (1900–02), p. 352. See also Szegö's paper, *Bemerkungen zu einem Satz von J. H. Grace über die Wurzeln algebraischer Gleichungen*, Mathematische Zeitschrift, vol. 13 (1922), p. 28, in which the "Faltungssatz" is, in a sense, the complement of the theorem proved above.

$$\psi \frac{d^r f}{dz^r} - (k-r+1)\frac{d\psi}{dz}\frac{d^{r-1}f}{dz^{r-1}} + \frac{(k-r+1)(k-r+2)}{2}\frac{d^2\psi}{dz^2}\frac{d^{r-2}f}{dz^{r-2}} + \cdots + (-1)^r \frac{(k-r+1)(k-r+2)\cdots k}{r!}\frac{d^r\psi}{dz^r} f = 0. \tag{7}$$

If every root of f is within a circle C, and every root of ψ is without C, then every root of (7) *is within C.*

NORTHWESTERN UNIVERSITY,
EVANSTON, ILL.

DETERMINATION OF ALL GENERAL HOMOGENEOUS POLYNOMIALS EXPRESSIBLE AS DETERMINANTS WHOSE ELEMENTS ARE HOMOGENEOUS POLYNOMIALS*

BY

H. S. EVERETT

1. **Introduction.** Dickson† has shown that every binary and every ternary form, every quaternary quadratic, and a sufficiently general quaternary cubic form can be expressed as a determinant with linear elements, and that this property does not hold for any other form which is the general one of its degree and number of variables. The present paper is a generalization from the case of determinants with linear elements to that of determinants with elements of higher degrees.

In the first part it is proved that if the number of variables is greater than three, no general form of degree rs is expressible determinantally, that is, as a determinant of order r with homogeneous elements of degree s, $s \geqq 2$. The case of quadratic elements is considered in § 2. The second part treats the determinantal expression of binary and ternary forms. In § 5 binary forms are expressed rationally as determinants. In § 6 a general procedure is set up for ternary forms and the remaining sections show that this procedure is successful for ternary quartic and sextic forms.

I. Forms with $n > 3$ are not expressible determinantally

2. Theorem 1. *When the number of terms in the general form of degree $2r$ in n variables ($n \geqq 2$) exceeds $[\frac{1}{2}n(n+1) - 2]r^2 + 2$, the form is not expressible as a determinant whose elements are quadratic forms.*

Let D be any r-rowed determinant with quadratic elements homogeneous in the n variables $x_1, \cdots, x_n$, which represents the general form of degree $2r$. Its matrix M may be written in the form

$$M = x_1^2 M_{11} + x_1 x_2 M_{12} + \cdots + x_1 x_n M_{1n} + x_2^2 M_{22} + \cdots + x_n^2 M_{nn},$$

where each M_{ij} is a matrix whose r^2 elements are constants. We proceed to simplify the matrix M by the method of matrix multiplication employed by Dickson.‡

* Presented to the Society, December 28, 1921.

† These Transactions, vol. 22 (1921), p. 167, hereinafter referred to as D.

‡ D, p. 168.

Since D represents the general form, x_1^{2r} is present, whence the determinant of M_{11} is not zero. Thus M_{11} has an inverse M_{11}^{-1} such that $M_{11}\,M_{11}^{-1}$ is the identity matrix I. The new matrix

$$N = MM_{11}^{-1} = x_1^2 I + x_1 x_2 N_{12} + \cdots + x_n^2 N_{nn} \quad (N_{ij} = M_{ij} M_{11}^{-1})$$

has as its determinant $D/|M_{11}|$ and has unity as the coefficient of x_1^{2r}. Next we choose a matrix B such that $B^{-1} N_{12} B$ shall have a canonical form P_{12}. Then our matrix reduces to

$$P = B^{-1} NB = x_1^2 I + x_1 x_2 P_{12} + \cdots + x_n^2 P_{nn} \quad (P_{ij} = B^{-1} N_{ij} B).$$

In contrast with the case of linear elements, we cannot show in our case of quadratic elements that the roots $\lambda_1, \cdots, \lambda_r$ of the characteristic equation of N_{12} are all distinct. Hence we permit P_{12} to have the most general form* with t distinct λ's, $t \leqq r$. Further normalization of P is effected by the matrix K commutative with P_{12} as defined by Dickson.† In $K^{-1} P_{13} K$ we can specialize $r - 1$ elements.

In case $t = r$ no further specialization is possible. We have then, if we count the factor removed from D and the r distinct λ's of P_{12},

$$1 + r + [\tfrac{1}{2} n(n+1) - 2] r^2 - (r - 1)$$

or $[\frac{1}{2} n(n+1) - 2] r^2 + 2$ parameters in the modified matrix of D. In case $t < r$ the true number of parameters is less than that given above because we then have $t < r$ distinct λ's, and because further normalization is possible since the matrix K is not the most general matrix commutative with P_{12}. Therefore, the number of parameters given above is the true maximum and our theorem follows.

3. Theorem 2. *When the number of terms in the general form of degree rs in n variables ($n \geqq 2$) exceeds*

$$S = \left[\frac{n(n+1)\cdots(n+s-1)}{s!} - 2 \right] r^2 + 2,$$

it is not expressible as a determinant whose elements are s-ic forms.

The argument in § 2 applies when D is a determinant of order r whose elements are forms of degree $s \geqq 2$ in n variables, except as to the number of terms in the matrix M. This number is $[n(n+1)\cdots(n+s-1)]/s!$.

4. Theorem 3. *If‡ $n > 3$, $s \geqq 2$, the general n-ary rs-ic form cannot be expressed as a determinant of order r whose elements are n-ary s-ic forms.*

The number of terms in the general form of degree rs in n variables is

$$\binom{rs+n-1}{rs} = \frac{(rs+1)\cdots(rs+n-1)}{(n-1)!}.$$

* D, p. 178; Dickson, *Linear Groups*, 1901, p. 223; Bôcher, *Higher Algebra*, p. 293.

† D, p. 179.

‡ For $s = 1$ see D, Theorem II, p. 171.

Then by § 3 the general form of degree rs in n variables ($n \geqq 2$) cannot be expressed as a determinant with s-ic elements if

$$(1) \qquad \frac{(rs+1)\cdots(rs+n-1)}{(n-1)!} > S.$$

For $n = 2$ and $n = 3$, respectively, (1) becomes $0 > (r-1)[(s-1)r-1]$ and $0 > (r-1)[(3s-2)r-2]$, which hold for no value of r when $s \geqq 2$.

Now (1) holds if

$$(2) \qquad \frac{(rs+1)(rs+2)(2s+3)\cdots(2s+n-1)}{(n-1)!} > S$$

holds, where the left member of (2) has been obtained from that of (1) by replacing r by 2 everywhere except in the first two factors. We consider this relation first for $s = 2$. If $n \geqq 7$, $r \geqq 2$, we can cancel from the numerator and denominator of the left member the factors $7, \cdots, (n-1)$ (which are absent if $n = 7$), and then (2) holds if

$$\frac{(2r+1)(2r+2)n(n+1)(n+2)(n+3)}{6!} \geqq \frac{1}{4}\left[\frac{n(n+1)}{2} - 2\right](2r+1)(2r+2),$$

since the right member exceeds the value $[\frac{1}{2}n(n+1)-2]r^2+2$ of S for $s = 2$. The last inequality holds if

$$n(n+1)(n+2)(n+3)/180 \geqq \tfrac{1}{2}n(n+1) - 2;$$

that is, if $n(n-7)(n+1)(n+12)+360 \geqq 0$, which is obviously true for $n \geqq 7$. For $n = 4$, $n = 5$, $n = 6$, respectively, (2) becomes

$$(2r-1)(2r-2)(2r-3) > 0, \qquad (r-1)(4r^3+24r^2-19r+6) > 0,$$

$$(r-1)(8r^4+68r^3+238r^2-107r+30) > 0,$$

each of which evidently holds for $r \geqq 2$. So (2) holds for $s = 2$, $r \geqq 2$, $n \geqq 4$.

We proceed now to show by induction on s that (2) and hence (1) holds for $s \geqq 2$, $r \geqq 2$, $n \geqq 4$. We assume that (2) holds for $r \geqq 2$, $n \geqq 4$, and any given value of s. We wish to show that (2) implies the corresponding inequality with s replaced by $s+1$, viz.,

$$(3) \qquad \frac{[(s+1)r+1][(s+1)r+2](2s+5)\cdots(2s+n+1)}{(n-1)!} > \left[\frac{n(n+1)\cdots(n+s)}{(s+1)!} - 2\right]r^2 + 2.$$

To derive (3) from (2) we must add to the right member of (2)

$$\left[\frac{n(n+1)\cdots(n+s)}{(s+1)!} - \frac{n(n+1)\cdots(n+s-1)}{s!}\right] r^2 = \frac{(n-1)n(n+1)\cdots(n+s-1)}{(s+1)!} r^2,$$

and to the left member

$$\frac{(2s+5)\cdots(2s+n-1)}{(n-1)!}(ar^2+br+c),$$

where

$$a = 4(n-1)s^3 + (n^2+9n-4)s^2 + 2(n^2+3n+1)s + n(n+1),$$
$$b = 12(n-2)s^2 + 3(n^2+5n-10)s + 3n(n+1),$$
$$c = 8(n-3)s + 2(n^2+n-12).$$

To establish the induction we need

$$(4) \qquad (2s+5)\cdots(2s+n-1)(ar^2+br+c)/(n-1) \geqq (s+2)\cdots(s+n-1)r^2.$$

Since the s-factors on the left are each greater than the corresponding s-factors on the right, and since $br + c > 0$ for $n \geqq 3$, (4) is implied by

$$(5) \qquad a/(n-1) \geqq (s+2)(s+3)(s+4).$$

Now (5) holds for $n \geqq 4$ if it holds for $n = 4$ since the coefficients of s are all positive for these values of n. But for $n = 4$, (5) becomes

$$9s^3 + 21s^2 - 20s - 52 \geqq 0,$$

which obviously holds for $s \geqq 2$. Our induction is therefore complete and Theorem 3 follows.

II. Binary and ternary forms expressed determinantally

5. **Binary forms.** Since the binary form of degree rs can be factored into rs linear factors, it is expressible as an r-rowed determinant each of whose diagonal elements is a product of s linear factors, and whose elements outside the main diagonal are zero. Dickson has expressed the binary form rationally as a determinant with linear elements.* Every binary form of degree rs can be expressed rationally as a determinant of order r with s-ic elements:

*D, p. 175. This determinant can be reduced by elementary transformations to the present one.

$$a_0 x^{rs} + a_1 x^{rs-1} y + \cdots + a_{rs-1} x y^{rs-1} + a_{rs} y^{rs} =$$

$$\begin{vmatrix}
[a_0 x^s + a_1 x^{s-1} y & + \cdots + a_{s-1} x y^{s-1} + a_s y^s] & y^s & 0 & 0 & \cdots & 0 \\
[- a_{s+1} x^{s-1} y & - \cdots - a_{2s-1} x y^{s-1} - a_{2s} y^s] & x^s & y^s & 0 & \cdots & 0 \\
[+ a_{2s+1} x^{s-1} y & + \cdots + a_{3s-1} x y^{s-1} + a_{3s} y^s] & 0 & x^s & y^s & \cdots & 0 \\
[- a_{3s+1} x^{s-1} y & - \cdots - a_{4s-1} x y^{s-1} - a_{4s} y^s] & 0 & 0 & x^s & \cdots & 0 \\
\cdot & \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
[(-1)^r a_{s(r-2)+1} x^{s-1} y & + \cdots + (-1)^r a_{s(r-1)} y^s] & 0 & 0 & 0 & \cdots & y^s \\
[(-1)^{r+1} a_{s(r-1)+1} x^{s-1} y & + \cdots + (-1)^{r+1} a_{rs} y^s] & 0 & 0 & 0 & \cdots & x^s
\end{vmatrix}.$$

6. **Ternary quartic form.** The general ternary quartic may be given the form

$$a^2 b^2 + c^2 d^2 + e^2 f^2 - 2abcd - 2abef - 2cdef,$$

where a and b, c and d, e and f, are three pairs of bitangents of a Steiner set.* This function equals the determinant

$$\begin{vmatrix} (ab - cd - ef) & 2cd \\ 2ef & (ab - cd - ef) \end{vmatrix}.$$

The determination of the bitangents depends, however, upon the solution of equations of high order. In the case of the general ternary quartic form with no repeated factor Dickson's determinant † of order $r = 4$ with linear elements can be reduced to a determinant of order $r = 2$ with elements of degree $s = 2$ by the following elementary transformations. Multiply the third row and divide the fourth column by $X_4 + c_{44} z$; multiply the second column and divide the first row by $X_1 + c_{11} z$; multiply the first column by $- z$ and add to the second column; multiply the fourth row by $- z$ and add to the third row; multiply the second row and divide the second column by z. The resulting determinant is

$$\begin{vmatrix} l_1 l_2 - c_{21} z^2 & z^2 \\ l_1 (c_{32} l_4 - c_{42} z) - z (c_{31} l_4 - c_{41} z) & l_3 l_4 - c_{43} z^2 \end{vmatrix},$$

where $l_i = X_i + c_{ii} z$. No elementary transformations have been found which accomplish a similar reduction of Dickson's determinant of order $rs > 4$.

7. **General ternary form.** The remainder of this paper is concerned with a method of expressing a sufficiently general ternary form f of degree rs as a determinant of order r whose elements are s-ic forms. We may assume that f is irreducible. For, if $f = f_1 f_2$ where f_1 and f_2 are of degree $r_1 s$ and $r_2 s$ respectively and expressible as determinants of order r_1 and r_2 whose matrices are M_1 and M_2, then f equals the determinant of the matrix

$$\begin{pmatrix} M_1 & O \\ O & M_2 \end{pmatrix},$$

* Miller, Blichfeldt, and Dickson, *Finite Groups*, 1916, p. 355.

† D, p. 174.

where O is a matrix all of whose elements are zero. With Dickson* we choose a triangle of reference such that, for $z = 0, f(x, y, 0) = X_1 X_2 \cdots X_{rs}$, where $X_i = x + \lambda_i y$ and $\lambda_1, \cdots, \lambda_{rs}$ are all distinct. Then

$$f = X_1 X_2 \cdots X_{rs} + \sum_{k=1}^{rs} z^k F_k(y, x),$$

where F_k is a binary form of degree $rs - k$. We shall attempt to express f as a determinant of the following special type:

$$D = \begin{vmatrix} X_1 \cdots X_s + z\phi_{11} & z\phi_{12} & \cdots & z\phi_{1r} \\ z\phi_{21} & X_{s+1} \cdots X_{2s} + z\phi_{22} & \cdots & z\phi_{2r} \\ \cdots & \cdots & \cdots & \cdots \\ z\phi_{r1} & z\phi_{r2} & \cdots & X_{rs-s+1} \cdots X_{rs} + z\phi_{rr} \end{vmatrix},$$

where each ϕ_{ij} is a ternary form of degree $s - 1$. There are $\frac{1}{2}(rs + 1)(rs + 2)$ coefficients in f. In identifying $f(x, y, 0)$ with $X_1, \cdots, X_{rs}, rs + 1$ of these coefficients have been fixed. So the identification of D with f involves $\frac{1}{2}rs(rs + 1)$ conditions on as many coefficients of the ϕ_{ij}. We therefore assign simple values to the remaining $\frac{1}{2}rs(r - 1)$ coefficients of the ϕ_{ij}; in each ϕ_{ij} $(j > i)$ we set the first s coefficients equal to zero with the exception of the first coefficient of each ϕ_{ij} $(j = i + 1)$ which we set equal to one.

We then expand D axially and consider the identification of the terms of this expansion, arranged in ascending powers of z, with the corresponding terms of any given form f. First, the terms linear in z will be identical with $zF_1(y, x)$ of f, where F_1 is any given binary form of order $rs - 1$, if they are equal for the rs values $x = -\lambda_i y$ $(i = 1, \cdots, rs)$. The resulting rs conditions involve the coefficients of the ϕ_{ij} of D. As a matter of fact, it will be seen in the following sections that they uniquely determine rs coefficients of the ϕ_{ii}. Next, the terms quadratic in z will be identical with $z^2 F_2(y, x)$ of f, where F_2 is of order $rs - 2$, if they are equal for† $x = -\lambda_i y$ $(i = 1, \cdots, rs - 1)$. The resulting $rs - 1$ equations involve coefficients of D. Proceeding in this manner to identify terms in z^k $(k = 1, \cdots, rs)$ we obtain a system of $rs + (rs - 1) + \cdots + 2 + 1 = \frac{1}{2}rs(rs + 1)$ equations in as many unknown coefficients of D. Since these equations involve arbitrary parameters $F_1, \cdots, F_{rs}$, and since by the implicit function theorem‡ solutions exist in the neighborhood of a point for which the jacobian of the equations does not vanish, the method outlined above effects the identification of D with f if the jacobian of these equations is not identically zero.

* D, p. 172.

† Or for any other choice of $rs - 1$ of the λ's.

‡ Bliss, *Princeton Colloquium Lectures*, 1913, p. 7. Attention is called in this connection to MacMillan's method of solution involving the introduction of a parameter and the expression of all solutions in power series of this parameter. See MacMillan, Mathematische Annalen, vol. 72, p. 180.

8. **Ternary quartic form.** A sufficiently general irreducible ternary quartic

$$f = X_1 X_2 X_3 X_4 + \sum_{i=1}^{4} z^i F_i(y, x), \qquad X_i = x + \lambda_i y,$$

with distinct λ's can be expressed as the determinant

$$D = \begin{vmatrix} X_1 X_2 + (h_{11} x + k_{11} y + l_{11} z) z & (x + l_{12} z) z \\ (h_{21} x + k_{21} y + l_{21} z) z & X_3 X_4 + (h_{22} x + k_{22} y + l_{22} z) z \end{vmatrix}.$$

The ten equations which arise in the identification of D with f follow. For brevity we write the frequently recurring expression $\lambda_u - \lambda_v$ as (uv), and $k_{uv} - \lambda_w h_{uv}$ as (uvw).

$$\begin{cases} (3i)(4i)(11i) = F_1(1, -\lambda_i) & (i = 1, 2), \\ (1j)(2j)(22j) = F_1(1, -\lambda_j) & (j = 3, 4); \end{cases}$$

$$\begin{cases} (3i)(4i) l_{11} + (11i)(22i) + \lambda_i (21i) = F_2(1, -\lambda_i) & (i = 1, 2), \\ (13)(23) l_{22} + (113)(223) + \lambda_3 (213) = F_2(1, -\lambda_3); \end{cases}$$

$$(11i) l_{22} + (22i) l_{11} - (21i) l_{12} + \lambda_i l_{21} = F_3(1, -\lambda_i) \qquad (i = 1, 2);$$

$$l_{11} l_{22} - l_{12} l_{21} = F_4.$$

The first four equations are linear and determine uniquely k_{11}, h_{11}, k_{22}, h_{22}, since their determinant is

$$M_4 = m_1 m_2 = (31)^2 (41)^2 (32)^2 (42)^2 (21) (43),$$

where

$$m_1 = (31)(41)(32)(42) \begin{vmatrix} 1 & -\lambda_1 \\ 1 & -\lambda_2 \end{vmatrix}, \qquad m_2 = (13)(23)(14)(24) \begin{vmatrix} 1 & -\lambda_3 \\ 1 & -\lambda_4 \end{vmatrix},$$

and is different from zero in view of the fact that the λ's are distinct. Since this system falls naturally into groups of 4, 3, 2, and 1 equations each, the Laplacian expansion of the jacobian of the system by minors of the last row, of the eighth and ninth rows, of the fifth, sixth, and seventh rows, and of the first four rows yields most readily the complete coefficient of the various terms. If the jacobian has its columns in the order k_{11}, h_{11}, k_{22}, h_{22}, k_{21}, h_{21}, l_{22}, l_{11}, l_{21}, l_{12}, we find on the main diagonal the only term of the expansion actually involving $l_{21} k_{22}$. Its coefficient is

$$M_4 M_3 M_2 M_1 = (21)^3 (31)^3 (32)^3 (41)^2 (42)^2 (43) \lambda_1 \lambda_2,$$

where

$$M_1 = -1, \qquad M_2 = \begin{vmatrix} 1 & \lambda_1 \\ 1 & \lambda_2 \end{vmatrix}, \qquad M_3 = \begin{vmatrix} \lambda_1 & -\lambda_1^2 & 0 \\ \lambda_2 & -\lambda_2^2 & 0 \\ \lambda_3 & -\lambda_3^2 & (13)(23) \end{vmatrix}.$$

We may assume that $\lambda_1 \lambda_2 \neq 0$, since if one of the four distinct λ's is zero it may be taken initially to be λ_4. Then since the λ's are distinct, the coefficient of $l_{21} k_{22}$ is different from zero and the jacobian does not vanish identically.

9. **Ternary sextic form, $s = 3$.** A sufficiently general irreducible ternary

sextic

$$f = X_1 X_2 \cdots X_6 + \sum_{k=1}^{6} z^k F_k(y, x), \qquad X_i = x + \lambda_i y, \tag{6}$$

with λ's distinct can be expressed as the determinant D with cubic elements

$$D = \begin{vmatrix} X_1 X_2 X_3 + z\phi_{11} & z\phi_{12} \\ z\phi_{21} & X_4 X_5 X_6 + z\phi_{22} \end{vmatrix},$$

where $\phi_{ij} = a_{ij} x^2 + b_{ij} xy + c_{ij} y^2 + d_{ij} xz + e_{ij} yz + f_{ij} z^2$, and where in particular $a_{12} = 1$, $b_{12} = c_{12} = 0$. The identification of D with f involves the following 21 equations. We use the abbreviated notations $(uv) = \lambda_u - \lambda_v$, $(uvw) = e_{uv} - \lambda_w d_{uv}$, and $[uvw] = a_{uv}\lambda_w^2 - b_{uv}\lambda_w + c_{uv}$.

$$\begin{cases} (4i)(5i)(6i)[11i] = F_1(1, -\lambda_i) & (i = 1,2,3), \\ (1j)(2j)(3j)[22j] = F_1(1, -\lambda_j) & (j = 4,5,6); \end{cases}$$

$$\begin{cases} (4i)(5i)(6i)(11i) + [11i][22i] - \lambda_i^2[21i] = F_2(1, -\lambda_i) & (i = 1,2,3), \\ (1j)(2j)(3j)(22j) + [11j][22j] - \lambda_j^2[21j] = F_2(1, -\lambda_j) & (j = 4,5); \end{cases}$$

$$\begin{cases} (4i)(5i)(6i)f_{11} + [11i](22i) + [22i](11i) \\ \qquad - \lambda_i^2(21i) - [21i](12i) = F_3(1, -\lambda_i) \quad (i = 1,2,3), \\ (14)(24)(34)f_{22} + [114](224) + [224](114) \\ \qquad - \lambda_4^2(214) - [214](124) = F_3(1, -\lambda_4); \end{cases}$$

$$[11i]f_{22} + [22i]f_{11} + (11i)(22i) - \lambda_i^2 f_{21} - [21i]f_{12} \\ - (12i)(21i) = F_4(1, -\lambda_i) \quad (i = 1,2,3);$$

$$(11i)f_{22} + (22i)f_{11} - (12i)f_{21} - (21i)f_{12} = F_5(1, -\lambda_i) \qquad (i = 1,2);$$

$$f_{11}f_{22} - f_{12}f_{21} = F_6.$$

With the columns of the jacobian arranged in the order $a_{11}, b_{11}, c_{11}, a_{22}, b_{22}, c_{22}, a_{21}, b_{21}, c_{21}, d_{22}, e_{22}, d_{21}, e_{21}, d_{11}, e_{11}, f_{11}, f_{22}, f_{21}, d_{12}, e_{12}, f_{12}$, we find on the main diagonal the complete term of the expansion involving $f_{21}^3 b_{11} c_{22}^3$. Its coefficient is

$$M_6 M_5 M_4 M_3 M_2 M_1 = \lambda_1^2\lambda_2^2\lambda_3^2 (12)^5 (13)^4 (14)^4 (15)^3 (16)^2 (23)^4 \\ \times (24)^4 (25)^3 (26)^2 (34)^4 (35)^3 (36)^2 (45)^2 (46)(65),$$

where

$$M_1 = -1, \qquad M_2 = \begin{vmatrix} \lambda_1 & -1 \\ \lambda_2 & -1 \end{vmatrix}, \qquad M_3 = \begin{vmatrix} 1 & -\lambda_1 & -\lambda_1^2 \\ 1 & -\lambda_2 & -\lambda_2^2 \\ 1 & -\lambda_3 & -\lambda_3^2 \end{vmatrix},$$

$$M_4 = \begin{vmatrix} \lambda_1^3 & -\lambda_1^2 & -\lambda_1 & 1 \\ \lambda_2^3 & -\lambda_2^2 & -\lambda_2 & 1 \\ \lambda_3^3 & -\lambda_3^2 & -\lambda_3 & 1 \\ \lambda_4^3 & -\lambda_4^2 & -\lambda_4 & 1 \end{vmatrix}, \qquad M_5 = \begin{vmatrix} -\lambda_1^4 & \lambda_1^3 & -\lambda_1^2 \\ -\lambda_2^4 & \lambda_2^3 & -\lambda_2^2 \\ -\lambda_3^4 & \lambda_3^3 & -\lambda_3^2 \end{vmatrix} \cdot \begin{vmatrix} -\lambda_4 & 1 \\ -\lambda_5 & 1 \end{vmatrix},$$

$$M_6 = \begin{vmatrix} \lambda_1^2 & -\lambda_1 & 1 \\ \lambda_2^2 & -\lambda_2 & 1 \\ \lambda_3^2 & -\lambda_3 & 1 \end{vmatrix} \cdot \begin{vmatrix} \lambda_4^2 & -\lambda_4 & 1 \\ \lambda_5^2 & -\lambda_5 & 1 \\ \lambda_6^2 & -\lambda_6 & 1 \end{vmatrix}.$$

We may assume that $\lambda_1\lambda_2\lambda_3 \neq 0$, since if one of the six distinct λ's is zero, it may be taken initially to be λ_6. Then since the λ's are distinct this coefficient is different from zero and the jacobian does not vanish identically.

10. **Ternary sextic form,** $s = 2$. A sufficiently general ternary sextic (6) can also be expressed as the determinant D with quadratic elements

$$D = \begin{vmatrix} X_1 X_2 + z\phi_{11} & z\phi_{12} & z\phi_{13} \\ z\phi_{21} & X_3 X_4 + z\phi_{22} & z\phi_{23} \\ z\phi_{31} & z\phi_{32} & X_5 X_6 + z\phi_{33} \end{vmatrix},$$

where $\phi_{ij} = h_{ij}\, x + k_{ij}\, y + l_{ij}\, z$ and $h_{12} = h_{23} = 1$, $h_{13} = k_{12} = k_{13} = k_{23} = 0$. Let Δ be the determinant $|l_{ij}|$ of third order, and let L_{ij} be the minor of l_{ij} in this determinant. In the following 21 equations which arise in the identification* of D with (6) we use the abbreviations

$$(uv) = \lambda_u - \lambda_v, \qquad (uvw) = k_{uv} - \lambda_w\, h_{uv}:$$

$$\begin{cases} (3i)(4i)(5i)(6i)(11i) = F_1(1, -\lambda_i) & (i = 1,2), \\ (1j)(2j)(5j)(6j)(22j) = F_1(1, -\lambda_j) & (j = 3,4), \\ (1k)(2k)(3k)(4k)(33k) = F_1(1, -\lambda_k) & (k = 5,6); \end{cases}$$

$$\begin{cases} (3i)(4i)(5i)(6i)l_{11} + (3i)(4i)(11i)(33i) \\ \quad + (5i)(6i)(11i)(22i) + (5i)(6i)\lambda_i(21i) = F_2(1, -\lambda_i) & (i = 1,2), \\ (13)(23)(53)(63)l_{22} + (13)(23)(223)(333) \\ \quad + (13)(23)\lambda_3(323) + (53)(63)(113)(223) \\ \quad + (53)(63)(213)\lambda_3 = F_2(1, -\lambda_3), \\ (1k)(2k)(3k)(4k)l_{33} + (1k)(2k)(22k)(33k) \\ \quad + (1k)(2k)\lambda_k(32k) + (3k)(4k)(11k)(33k) \\ \quad = F_2(1, -\lambda_k) & (k = 5,6); \end{cases}$$

$$\begin{cases} (3i)(4i)[(11i)l_{33} - (31i)l_{13} + (33i)l_{11}] \\ \quad + (5i)(6i)[(11i)l_{22} - (21i)l_{12} + (22i)l_{11} + \lambda_i l_{21}] \\ \quad + (11i)(22i)(33i) + (11i)(32i)\lambda_i \\ \quad + (21i)(33i)\lambda_i + (31i)\lambda_i^2 = F_3(1, -\lambda_i) & (i = 1,2), \\ (1k)(2k)[(22k)l_{33} - (32k)l_{23} + (33k)l_{22} + \lambda_k l_{32}] \\ \quad + (3k)(4k)[(11k)l_{33} - (31k)l_{13} + (33k)l_{11}] \\ \quad + (11k)(22k)(33k) + (11k)(32k)\lambda_k \\ \quad + (21k)(33k)\lambda_k + (31k)\lambda_k^2 = F_3(1, -\lambda_k) & (k = 5,6); \end{cases}$$

* See second footnote, § 7. Here it was found advantageous to select the λ's in the order $\lambda_1, \lambda_2, \lambda_5, \lambda_6, \lambda_3, \lambda_4$, since this selection gives the minimum number of terms in certain equations. Thus, the tenth and eleventh equations contain fewer terms than does the ninth equation.

$$\left\{\begin{aligned}&(5i)(6i)L_{33}+(3i)(4i)L_{22}+l_{13}[(21i)(32i)-(22i)(31i)]\\&\quad-l_{23}[(11i)(32i)+\lambda_i(31i)]\\&\quad+l_{33}[(11i)(22i)+\lambda_i(21i)]\\&\quad+\lambda_i[(11i)l_{22}-(31i)l_{12}+(32i)l_{11}+\lambda_i l_{31}]\\&\quad+(33i)[(11i)l_{22}-(21i)l_{12}+(22i)l_{11}+\lambda_i l_{21}]\\&\quad=F_4(1,-\lambda_i)\qquad(i=1,2);\\&\text{Similarly for } i=6 \text{ with first term replaced by } (16)(26)L_{11};\end{aligned}\right.$$

$$(11i)L_{11}-(21i)L_{21}+(31i)L_{31}+\lambda_i L_{12}+(22i)L_{22}$$
$$-(32i)L_{32}+\lambda_i L_{23}+(33i)L_{33}=F_5(1,-\lambda_i)\quad(i=1,2);$$

$\Delta = F_6$.

With the columns of the jacobian arranged in the order k_{11}, h_{11}, k_{22}, h_{22}, k_{33}, h_{33}, k_{21}, h_{21}, l_{22}, k_{32}, h_{32}, k_{31}, h_{31}, l_{33}, l_{32}, l_{11}, l_{21}, l_{31}, l_{12}, l_{23}, l_{13}, we find on the main diagonal the complete term of the expansion involving $l_{21}^2\, l_{32}^2\, h_{11}\, k_{22}^2\, k_{33}^3$. Its coefficient is

$$M_6\, M_5\, M_4\, M_3\, M_2\, M_1 = \lambda_1^2\lambda_3^3\lambda_5\lambda_6\,(12)^5(13)^3(14)^2(15)^5(16)^6$$
$$\times(23)^3(24)^2(25)^5(26)^6(34)(35)^3(36)^3(45)^2(46)^2(56)^3,$$

where

$$M_1=1,\qquad M_2=-\begin{vmatrix}1&\lambda_1\\1&\lambda_2\end{vmatrix},\qquad M_3=\begin{vmatrix}1&\lambda_1&\lambda_1^2\\1&\lambda_2&\lambda_2^2\\1&\lambda_6&\lambda_6^2\end{vmatrix},$$

$$M_4=\begin{vmatrix}\lambda_1^2&-\lambda_1^3&0&0\\\lambda_2^2&-\lambda_2^3&0&0\\\lambda_5^2&-\lambda_5^3&(15)(25)&(15)(25)\lambda_5\\\lambda_6^2&-\lambda_6^3&(16)(26)&(16)(26)\lambda_6\end{vmatrix},$$

$$M_5=(13)(23)(53)(63)(51)^2(61)^2(52)^2(62)^2\lambda_1\lambda_2\lambda_5\lambda_6$$
$$\times\begin{vmatrix}1&-\lambda_1\\1&-\lambda_2\end{vmatrix}\cdot\begin{vmatrix}1&-\lambda_5\\1&-\lambda_6\end{vmatrix},$$

$$M_6=(31)^2(41)^2(51)^2(61)^2(32)^2(42)^2(52)^2(62)^2(53)^2$$
$$\times(63)^2(54)^2(64)^2\begin{vmatrix}1&-\lambda_1\\1&-\lambda_2\end{vmatrix}\cdot\begin{vmatrix}1&-\lambda_3\\1&-\lambda_4\end{vmatrix}\cdot\begin{vmatrix}1&-\lambda_5\\1&-\lambda_6\end{vmatrix}.$$

We may assume that $\lambda_1\lambda_2\lambda_5\lambda_6 \neq 0$ since if one of the six distinct λ's is zero it may be taken initially to be λ_4. Then since the λ's are distinct the coefficient of this term is different from zero and the jacobian does not vanish identically.

No attempt has been made to prove that for ternary forms higher than the sextic the corresponding jacobian is not identically zero.

Bucknell University,
Lewisburg, Pa.

GENERAL VECTOR CALCULUS *

BY

JAMES BYRNIE SHAW

Introduction

This paper presents results from various papers read before the Society during a period of several years. These are indicated in the footnotes. The calculus is independent of the number of dimensions of the space in which the vectors are supposed to be placed. Indeed the vectors are for the most part supposed to be imbedded in a space of an infinity of dimensions, this infinity being denumerable sometimes but more often non-denumerable. In the sense in which the term is used every vector is an infinite vector as regards its dimensionality. The reader may always make the development concrete by thinking of a vector as a function of one or more variables, usually one, and involving a parameter whose values determine the "dimensionality" of the space. The values the parameter can assume constitute its "spectrum." It must be emphasized however that no one concrete representation is all that is meant, for the vector is in reality an abstract entity given by its definition, that is to say, postulationally. The case is analogous to that of "group" in which the "operators" are generally not operators at all, since they have nothing to operate upon, but are abstract entities, defined by postulates. Always to interpret vectors as directed line-segments or as expansions of functions is to limit the generality of the subject to no purpose, and actually to interfere with some of the processes. It is sufficient to notice that in any case the theorems may be tested out in any concrete representation. The author desires to state that he discriminates between the terms *function, values of a function, expansion of a function*. The function is the law which enables one to ascertain the values. The table of values is not the function. The expansion is the system of coefficients of the basis functions by means of which we may have a particular representation of the function. Further a system of coefficients does not constitute a *vector*; such a system the author calls a *multiplex*.

It is true that a multiplex appeals to many mathematicians as a sufficient entity to serve as vector, but a multiplex does not specify a vector till the basis (or in geometry the set of axes) is specified. A system of coefficients does not define the same function when attached to powers of x as when attached to sines in a Fourier series; the vector $(1, 2, 3)$ has no meaning till the directions for the

* Presented to the Society at various meetings, 1909–1921.

Trans. Am. Math. Soc. **14.**

numbers are assigned; and $1 - x^2$ is the same function whether expanded in a Fourier series or a Bessel series, though its multiplex is different. The essence indeed of a vector calculus lies in its ability to produce invariantive and covariantive expressions, and it does this because it uses vectors directly and not some particular representation of the vectors. A calculus which produces formulas in x, y, z, even though they may be at once replaced by x', y', z', and preserve the same formulas, may be an invariant calculus, but it is not a vector calculus. At the best, symbols which stand for such expressions referred to axes for their meaning are merely short-hand symbols.

The method developed here has been found by years of use to be flexible, smooth-running, simple, and to include all the systems of vectors now extant. Its expressions are all covariantive for the simple reason they have no dependence upon any system of reference. They are compact and readily interpreted in such applications as differential geometry. It furnishes the basis for integral equations, and the general theory of distributive operators. In the abstract the intention is to create a calculus of symbols which are defined by the rules of the calculus. Any entities that submit to these rules may be represented by the symbols so far as this calculus is concerned, though from other points of view they must be treated differently. These symbols stand for entities which are subject to certain *processes*. A process is not a function; it produces entities which are functions of the entities subject to the process, but the process is not the function. For instance we use a process called *addition*, and one called *multiplication*. There are several called "A" processes which are not "multiplications." The author is aware that some mathematicians like to call some of these A processes by such names as *inner multiplication, outer multiplication, indeterminate multiplication*, and the like, but he has not yet observed that this multiplicity of names has been any advantage in the development of the subject.

On the side of vectors in finite space (as to dimensions) this method includes the Clifford algebras,* and might be supposed to be identical with the general treatment of Clifford. The question has been asked: Why not assume a system of *units*, $i_1, i_2, \cdots, i_n$, whose squares are -1, and such that $i_p i_q = -i_q i_p$ when $p \neq q$? All the formulas in the finite cases may be worked out from this point of departure. They might then be generalized to the case of a denumerable infinity of units. This would indeed at least simplify much of the extant work on the uses of orthogonal functions, vectors in a "Hilbert space," etc. But when we take up the most general case of a non-denumerable space it would fail. The present development has no reference to any dimensions.

* Consult Clifford's *Works;* Joly, Proceedings of the Royal Irish Academy (3), vol. 5 (1897), pp. 73–123; vol. 6 (1900), pp. 13–18; M'Aulay, Proceedings of the Royal Society of Edinburgh, vol. 28 (1908), pp. 503–585.

This must be well understood by the reader, for it is fundamental. The Grassmann treatments are hampered by their assumption that the "outer product" of all the basis units of the system is a numerical value. Without knowing the number of dimensions of the space the formulas cannot always be written. The formulas here are equally valid whether the vectors lie in a space of three, three thousand, a denumerable number, or a non-denumerable number of dimensions. There is no basis, for every vector is an entity unique in itself. It defines with its numerical multiples a linear space, any two define with their numerical multiples and under addition a binary space, any n an n-ary space,—under certain restrictions to be noted.

I. FUNDAMENTALS

1. Definitions

1. **Vectors.** Vectors are represented in notation throughout by small Greek letters, or by expressions built up with small Greek letters, and other symbols. Capital Greek letters invariably mean operators. Roman (italic) letters mean numerical values selected from a given domain, or field. An exception sometimes is furnished in using π in the ordinary sense.

Vectors are defined by stating the conditions to which they are subject. These are as follows.

(1) Each vector gives rise to a set of vectors, called its *multiples*, represented by writing the symbol of the vector and a symbol for a number (or mark) chosen from the given field, thus from α we have $a\alpha$, 3α, $\pi\alpha$, and the like. No distinction is made between $a\alpha$ and αa. The multiples of a vector are as numerous as the marks or numbers of the field. Any one of the multiples may itself be taken as the initial vector and the others, including the original vector, are then multiples of it. No vector is a unit vector absolutely, but only relatively.

(2) In case we have one more vector which is not among the multiples of the vector α, say β, then the multiples of α and the multiples of β enable us to form a *binary complex of vectors*, indicated by $x\alpha + y\beta$, where x and y independently assume all values in the field. The vector $\xi = x\alpha + y\beta$ is a unique vector, but is said to be dependent upon α and β, or not to be linearly independent of them. We call the process of forming these vectors of the complex *addition of vectors*. We assume as a matter of definition that

$$\xi = x\alpha + y\beta = y\beta + x\alpha,$$

and further that, in the case of still a third vector, not in the complex of α, and β, say γ, if $\xi = x\alpha + y\beta$, $\zeta = y\beta + z\gamma$, $\sigma = \xi + z\gamma$, $\tau = x\alpha + \zeta$, then in all cases $\sigma = \tau$.

This is called *associativity of addition*.

(3) If $\xi = \sum x_n \alpha_n$, then the multiples of ξ, as $c\xi$, are also to be formed from the same multiples of $x_1 \alpha_1, x_2 \alpha_2, \cdots$, that is

$$c\xi = \sum cx_n \alpha_n .$$

We see that by this last condition

$$x\alpha + y\alpha + \cdots = (x + y + \cdots)\alpha .$$

(4) If 0 is included in the field, then we have vectors $0\alpha, 0\beta, \cdots$ and these are indistinguishable.

It follows that we may consider $\alpha + 0\xi = \alpha$ in all cases. Further we see that the multiples of α, β, γ, etc., which determine a complex are to be considered as in the complex. Hence the vectors which determine a complex are not unique. Whatever vectors are used to determine a complex are called a *basis* for the complex in case they constitute an *irreducible* set, in the sense that no one of them is included in the complex of the others. That is the same as saying that $\alpha_1, \alpha_2, \cdots, \alpha_n$ form a basis for a complex in case

$$0\rho = x_1 \alpha_1 + x_2 \alpha_2 + \cdots + x_n \alpha_n$$

only for $x_1 = x_2 = \cdots = x_n = 0$. They are said then to be linearly independent.

(5) When the number of linearly independent vectors is not finite we indicate them by using a parameter attached to the vector symbol, sometimes as a subscript, as in the case of a denumerable infinity, thus $\alpha_1, \alpha_2, \cdots, \alpha_\omega$; and sometimes as a parameter in a parenthesis, as $\alpha(i)$, as in the case of a non-denumerable infinity. The complex in the first case is indicated by

$$\xi = \sum_1^\omega x_n \alpha_n \qquad \text{or} \qquad \sum x_n \alpha(n) .$$

The complex in the second case is indicated by

$$\xi = \int_a^b x(i)\alpha(i)\, di .$$

As an instance of each case, so far as our conditions go we may consider that functions of a single variable are vectors. We would then have, using z for the variable, i for the parameter, in the first case, an instance from expansions in orthogonal functions,

$$\xi(z) = \sum a_n \phi_n(z);$$

in the second case, an instance from the expression of a function as a definite integral,

$$\xi(z) = \int_a^b a(i)\phi(i, z)\, di .$$

When we apply vectors to functions we need to note there are further restrictions on account of the problems of analysis. We are concerned here largely with the problems of algebra only.

When we undertake to find the coefficients (from the field) which would enable us to express a given vector as belonging to the complex of a given basis (*finding the components of the vector*), in case it is possible to express the vector thus, that is, in case the vector does belong to the complex, we say we *resolve the given vector* on the given basis. This is exemplified in the process of finding components of velocity, or force, or vector field, or Fourier coefficients, or solving a linear homogeneous integral equation. The conditions, in the function case again, that such resolution may take place, have been extensively studied in certain special cases. The study of such conditions as result from these analytic problems has been carried far by E. H. Moore, whose papers on General Analysis may be referred to. The definitions of norm, Hermitian square, generalized theorem of Pythagoras, etc., find their place in such investigations.

(6) Vectors are further subject to conditions which are the main object of the present paper. These result from the formation of *vectors of grade* $2, 3, \cdots$, and their combinations into complexes, and combinations into complexes of mixed grades. These will be treated separately under the heads: *the accretive process, the decretive process, the Hamiltonian process.* The latter will also be called *the product.*

2. The accretive process

2. **Alternants.*** In general an expression is alternant when it changes sign with the interchange of two of its elements. The usage here is not contradictory to this but supplements it considerably. As an example of an alternant let us consider n functions of n different variables, giving us n^2 different functions

$$\alpha_i(s_j) \qquad (i, j = 1, 2, \cdots, n).$$

Then the expression $\sum \pm \alpha_{i_1}(s_1)\alpha_{i_2}(s_2)\cdots\alpha_{i_n}(s_n)$, where the subscripts i are the numbers $1, 2, \cdots, n$ in some permutation, all permutations occurring, and the sign is $+$ or $-$ according to the number of inversions, is an expression which changes sign if any two subscripts are interchanged. It is therefore an alternant.

The alternants we introduce are defined by the following conditions:

(1) From any n vectors we may construct an alternant, written $A_n \cdot \alpha_1\alpha_2\cdots\alpha_n$, called a simple vector of grade n.

* The remainder of part I contains the results of papers read before the Society as follows: Chicago Section, December 26, 1913, *The two fundamental operations of general vector analysis*; December 31, 1909, *On Hamiltonian products*; Southwestern Section, November 27, 1909, *Scalars of lineal products.*

(2) The interchange of any two vectors changes the sign.

(3) If any one of the vectors is in a complex, say $\alpha_i = \sum a_j \beta_j$ where the $\sum$ may be used also to include the non-denumerable case

$$\alpha(i) = \int a(j)\beta(j)\,dj,$$

then

$$A_n \cdot \alpha_1 \alpha_2 \cdots \alpha_i \cdots \alpha_n = \sum a_j A_n \cdot \alpha_1 \alpha_2 \cdots \beta_j \cdots \alpha_n.$$

It follows at once that if any two vectors are equal the alternant vanishes, that if any vector is in the complex of the others the alternant vanishes, and that to any vector may be added any vector in the complex of the others without altering the alternant. The distributivity property (3) is very important.

(4) $A_1 \cdot \alpha = \alpha$. The alternants begin with A_2. We preserve the notation since it is valid in general theorems.

(5) If $A_n \cdot \alpha_1 \alpha_2 \cdots \alpha_n = 0$ then for some set of values $x_1, x_2, \cdots, x_n$ not all zero

$$x_1 \alpha_1 + x_2 \alpha_2 + \cdots + x_n \alpha_n = 0.$$

This makes either equation a necessary and sufficient condition for the other. It precludes using as the alternant defined above *any* alternant expression. For instance in quaternions we might use as $A_2 \cdot \alpha\beta$ the expression $V \cdot \theta(\alpha)\theta(\beta)$ where $\theta(\ )$ is a linear vector function, but it would be necessary to choose θ as non-singular. Otherwise the expression might vanish because α or β happened to be an invariant line for the root 0. Similarly we might take as $A_2 \cdot \alpha\beta$ the expression

$$\int_a^b \int_a^t [\alpha(s)\beta(t) - \beta(s)\alpha(t)]\phi(s)\,ds dt,$$

but we should have to place restrictions upon what $\phi(\ )$ could be. Indeed in either example we can consider that we do not really have the simple alternant, but a function of the alternant, for they may be written respectively $\Phi V \cdot \alpha\beta$, and $\Phi[\alpha(s)\beta(t) - \beta(s)\alpha(t)]$. The function might vanish when the elementary alternant does not. We limit the meaning of $A_n \cdot$ otherwise, however, only by the postulates.

3. **Vector of grade n.*** A basis of alternants of grade n gives a complex of grade n, and any member of the complex is a vector of grade n. A vector of grade n is indicated by V_n. It does not have to reduce to a single alternant, as for instance $A_2 \cdot \alpha\beta + A_2 \cdot \gamma\delta$ may or may not reduce to the form $A_2 \cdot \epsilon\zeta$. Many vectors of grade n do however reduce to single terms which are represented by the symbols A_n. Such is the case to be defined, which is written

*This is not the "polyadic" of some authors.

$A_n \cdot \alpha_1 \alpha_2 \cdots \alpha_{n+k}$ where $k \neq 0$. Such expressions have properties to be proved which may be stated in terms of the permutation groups under which they are invariant. In case $k = 0$ we might say that $A_n \cdot \alpha_1 \alpha_2 \cdots \alpha_n$ is invariant under the alternating group of permutations of its vectors, and changes sign for the remaining substitutions of the symmetric group. Other expressions constructed from an initial term by applying the substitutions of some group would be interesting and useful, but little study of them has occurred.

4. **Accretive process.** The accretive process is now defined by the equation

$$A_{p+q} \cdot A_p \cdot \alpha_1 \alpha_2 \cdots \alpha_p A_q \cdot \beta_1 \beta_2 \cdots \beta_q = A_{p+q} \cdot \alpha_1 \alpha_2 \cdots \alpha_p \beta_1 \beta_2 \cdots \beta_q ,$$

where p, q, are any integers. Hence if the alternants are interchanged the sign is multiplied by $(-1)^{pq}$.

3. The decretive process

5. **Scalars.** The term scalar was first applied by Hamilton to quaternions which could be included in the ordinary number scale. It is used here in the same sense but with the further application to a process which gives numbers (marks) belonging to the field, from vectors, whatever the grade of the vectors. We shall use $A_0 \cdot$ to indicate such results. We define first:

(1) $A_0 \cdot A_n \cdot \alpha_1 \alpha_2 \cdots \alpha_n A_n \cdot \beta_1 \beta_2 \cdots \beta_n$ is a number or mark of the coefficient field.

(2) $A_0 \cdot A_n \alpha_1 \alpha_2 \cdots \alpha_n A_n \cdot \beta_1 \beta_2 \cdots \beta_n$
$= A_0 \cdot A_n \cdot \beta_1 \beta_2 \cdots \beta_n A_n \cdot \alpha_1 \alpha_2 \cdots \alpha_n .$

This will be found to be a special case of a more general formula. In particular

$$A_0 \cdot \alpha\beta = A_0 \cdot \beta\alpha .$$

(3) $A_0 \cdot A_p ( ) A_q ( ) = 0, \quad p \neq q .$

(4) $A_0 \cdot (P + Q + \cdots)(R + S + \cdots)$
$= A_0 \cdot PR + A_0 \cdot PS + \cdots + A_0 \cdot QR + A_0 \cdot QS + \cdots$

for any expressions $P, Q, \cdots, R, S, \cdots$. That is, A_0 is distributive.

Instances of the scalar process are found in the expression called *inner product* by some mathematicians. In geometric vectors it is usually defined to be the product of the lengths of the vectors by the cosine of their angle, which angle is defined for the space involved. In functions we have

$$A_0 \cdot \alpha\beta = \int_a^b \alpha(s)\,\beta(s)\,ds .$$

We do not identify the scalar process A_0 with the scalar process in quaternions, though the latter is included. For we might take $A_0 \cdot \alpha\beta = S\alpha\theta(\beta)$ where

$\theta()$ is a self-transverse (symmetric, self-conjugate) linear vector operator which has no zero roots. Such a form is implied in the integration with a matrix defined by E. H. Moore in his General Analysis.

Scalars of vectors of grades higher than one are expressible by scalars of vectors of grade one.

6. **Orthogonality.** If the scalar of two vectors of any grade vanishes the two vectors are said to be *orthogonal*. A system of vectors may be mutually orthogonal, as for instance a system of mutually perpendicular vectors in space, or a system of functions such as $\alpha_i = \sqrt{2}\sin i\pi s$, under the definition

$$A_0 \cdot \alpha_i \alpha_j = \int_0^1 \alpha_i(s)\alpha_j(s)\,ds.$$

Biorthogonality. Two systems of vectors $\alpha(i)$, $\beta(j)$ are mutually biorthogonal in case we have

$$A_0 \cdot \alpha(i)\beta(j) = 0, \qquad i \neq j,$$
$$A_0 \cdot \alpha(i)\beta(i) \neq 0.$$

Biorthogonality and orthogonality play an important part in the problem of expansions mentioned earlier. For instance they enable us to find the coefficients under proper restrictions.

7. **Decretive process.** We are now in a position to define the general decretive process. We define first the expression

$$A_{n-1} \cdot \alpha A_n \cdot \beta_1\beta_2\cdots\beta_n = \sum(-1)^{i-1} A_0 \cdot \alpha\beta_i \cdot A_{n-1} \cdot \beta_1 \underset{i}{\cdots} \beta_n$$

where the i written under $\beta_1\cdots\beta_n$ means that β_i is missing from the expression. That is $\beta_1 \underset{i}{\cdots} \beta_n$ means $\beta_1\cdots\beta_{i-1}\beta_{i+1}\cdots\beta_n$, with the particular cases $\beta_1 \underset{1}{\cdots} \beta_n$ means $\beta_2\cdots\beta_n$ and $\beta_1 \underset{n}{\cdots} \beta_n$ means $\beta_1\cdots\beta_{n-1}$. This expression is evidently a vector of grade $n-1$. The decretive process reduces the grade. We define further

$$A_{n-1} \cdot (A_n \cdot \beta_1\cdots\beta_n)\alpha = (-1)^{n-1} A_{n-1} \cdot \alpha A_n \cdot \beta_1\cdots\beta_n.$$

We shall see later that we are here following a general law of permutations which is the same as the determinant law for change of sign, but we have to notice that in A_0 the two vectors may be permuted without change of sign.

If we iterate the process just defined we would have

$$A_{n-2} \cdot \alpha_1 A_{n-1} \cdot \alpha_2 A_n \cdot \beta_1\cdots\beta_n$$
$$= \sum(-1)^{i+j} \begin{vmatrix} A_0 \cdot \alpha_1\beta_i & A_0 \cdot \alpha_1\beta_j \\ A_0 \cdot \alpha_2\beta_i & A_0 \cdot \alpha_2\beta_j \end{vmatrix} A_{n-2} \cdot \beta_1 \underset{i,j}{\cdots} \beta_n, \qquad i < j.$$

This is evident since the multiplier of $A_{n-2} \cdot \beta_1 \underset{i,j}{\cdots} \beta_n$ will come from two

sources; being, if we omit the signs, one way $A_0 \cdot \alpha_2 \beta_j \cdot A_0 \cdot \alpha_1 \beta_i$, and the other way $A_0 \cdot \alpha_2 \beta_i A_0 \cdot \alpha_1 \beta_j$. But if $i < j$ the sign for the first is $(-)^{i+j+2}$ and for the second $(-)^{i+j+3}$, hence they have opposite signs. It is easy to see by induction that for r iterations we have

$$A_{n-r} \cdot \alpha_1 A_{n-r+1} \cdot \alpha_2 \cdots A_{n-1} \cdot \alpha_r A_n \cdot \beta_1 \cdots \beta_n = \sum_{i_1, \cdots, i_r} (-)^{i_1 + i_2 + \cdots + i_r + r} |A_0 \cdot \alpha_k \beta_m| A_{n-r} \cdot \beta_1 \cdot \cdots \cdot \beta_n$$

where $k = 1, 2, \cdots, r$, $m = i_1, i_2, \cdots, i_r$, $i_1 < i_2 < \cdots < i_r$.

The formula for operation on the right instead of the left is easily written.

As an instance in functions, we may define $A_n \cdot \beta_1 \cdots \beta_n$ to be the determinant

$$\frac{1}{n!} \sum \pm \beta_1 (s_{h_1}) \beta_2 (s_{h_2}) \cdots \beta_n (s_{h_n}),$$

a function of n independent variables. If now we multiply by $\alpha_1 (s)$, we may produce the decretive process by forming the sum, with proper sign, of the integrals we have, when we set in turn $s = s_1$, $s = s_2$, $\cdots$, $s = s_n$.

Since the vectors α enter the expressions above in an alternating manner we are warranted in the final definition, for the decretive process on two vectors of any grades,

$$A_{n-r} \cdot A_r \cdot \alpha_1 \cdots \alpha_r A_n \cdot \beta_1 \cdots \beta_n = A_{n-r} \cdot \alpha_1 A_{n-r+1} \cdot \alpha_2 \cdots A_{n-1} \cdot \alpha_r A_n \cdot \beta_1 \cdots \beta_n, \qquad r \leqq n,$$

and we can show that

$$A_{n-r} \cdot A_n \cdot ( ) A_r \cdot ( ) = (-)^{r(n-1)} A_{n-r} \cdot A_r \cdot ( ) A_n \cdot ( ) \quad \text{when} \quad r \leqq n.$$

In either case if $r = n$ we have the important formula

$$A_0 \cdot A_n \cdot \alpha_1 \cdots \alpha_n A_n \cdot \beta_1 \cdots \beta_n = (-)^{\frac{1}{2}n(n-1)} |A_0 \cdot \alpha_i \beta_j| \qquad (i, j = 1, 2, \cdots, n).$$

The expression just found is sometimes called the Grammian when the α's and β's represent functions. In this case we define

$$A_0 \cdot \alpha\beta = \int_a^b \alpha (s) \beta (s) \, ds.$$

For n functions the vanishing of the Grammian is a necessary and sufficient condition of their being linearly dependent. However, this is really because of the presence of the $A_n \cdot$ of the n functions. It is really what vanishes. It is evident now that we may rewrite the definition of the general decretive process in a theorem

$$A_{n-r} \cdot A_r \cdot \alpha_1 \cdots \alpha_r A_n \cdot \beta_1 \cdots \beta_n = \sum_{i_1, \cdots, i_r} (-)^{\Sigma i + \frac{1}{2} r (r-1)} A_0 \cdot A_r \cdot \alpha_1 \cdots \alpha_r A_r \cdot \beta_{i_1} \cdots \beta_{i_r} A_{n-r} \beta_1 \cdot \cdots \cdot \beta_n.$$

It is to be observed that the sign is to be obtained by the determinant rule, since we pass β_{i_1} over $i_1 - 1$ preceding β's, β_{i_2} then over $i_2 - 2$ preceding β's, etc. The total inversions are then $i_1 + i_2 + \cdots + i_r + \frac{1}{2}r(r+1)$. This theorem is a particular case of a more general one to follow, which, however, is to be stated as a definition of both the accretive and the decretive processes for all cases.

The expression $A_0 \cdot A_n(\,)\,A_n(\,)$ is called the *norm* or *square of the intensity* of the vector A_n, as for instance $A_0 \cdot A_2 \cdot \alpha\beta . A_2 \cdot \alpha\beta$ is the square of the intensity of $A_2 \cdot \alpha\beta$. An alternant divided by its intensity, if this is not 0, gives a *unit alternant*.

By means of the two processes it is possible to construct the whole of a general vector calculus.*

8. When $A_0 \cdot A_n \cdot \alpha_1 \cdots \alpha_n A_n \cdot \beta_1 \cdots \beta_n = 0$, we know from the properties of determinants that if the rank is h, by proper addition of rows and columns we may reduce the determinant to one in which either $n - h$ rows or $n - h$ columns contain only 0, which in the present determinant will give vectors in these rows or columns that are in the complex of the vectors α or the complex of the vectors β, respectively. It follows that $n - h$ vectors in the complex of β's may be chosen so as to all be orthogonal to every vector α, or we may choose $n - h$ vectors in the complex of the α's which are all orthogonal to all the vectors β. It follows that the complex of the remaining h linearly independent vectors in the complex of the α's, and the remaining h linearly independent vectors in the complex of the β's, define two complexes of h dimensions, which have a common *incident* complex of h dimensions. We must not assume that the complex of order h of the vectors α can be identified with that of order h of the vectors β, since each vector of the first may be of the form $\sum_1^h x_i \gamma_i + \zeta_j$, $j = 1, \cdots, h$, and the vectors of the other complex of the form $\sum_1^h y_i \gamma_i + \xi_j$, $j = 1, \cdots, h$, where the vectors ζ and ξ are orthogonal to all the vectors γ, and all vectors ζ are orthogonal to all vectors ξ. The A_0 process leaves only the bilinear homogeneous forms $\sum x_i y_i$. It follows that the h vectors α are not necessarily expressible as in the complex of the h vectors β, and vice versa. We may say then that the complex of the n vectors which determine a vector of grade n, and the complex of the n vectors which determine another vector of grade n, may have with regard to each other a *rank* h, which is the order of two sub-complexes, one from each complex, which have a common sub-complex of order h. In the geometric case it is easy to see the significance of these facts. In the function spaces, however, they are also significant. When the general decretive expression vanishes, each scalar coefficient vanishes provided the vectors β are linearly independent. Hence all vectors α and β become subject to the remarks just made.

* This was presented to the Society in a paper read March 26, 1921, *On Hamiltonian products*, second paper. The contents are incorporated in this memoir.

9. In the alternant $A_n \cdot \alpha_1 \cdots \alpha_n$ we may substitute other vectors without changing the value of the alternant. These may be chosen so as to form a mutually orthogonal system if their intensities do not vanish. For it is evident that

$$\alpha_2 - \alpha_1 \frac{A_0 \cdot \alpha_1 \alpha_2}{A_0 \cdot \alpha_1 \alpha_1}$$

is orthogonal to α_1 and its substitution for α_2 will not change the value of the alternant. We may now find a vector orthogonal to both α_1, α_2 to substitute for α_3, etc. In short if we set

$$\begin{aligned}
&\alpha_1' = \alpha_1, \\
&\alpha_2' = A_1 \cdot \alpha_1 A_2 \cdot \alpha_1 \alpha_2 / A_0 \cdot \alpha_1 \alpha_1, \\
&\alpha_3' = A_1 \cdot A_2 \cdot \alpha_1 \alpha_2 A_3 \cdot \alpha_1 \alpha_2 \alpha_3 / A_0 \cdot A_2 \cdot \alpha_1 \alpha_2 A_2 \cdot \alpha_1 \alpha_2, \\
&\quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad .
\end{aligned}$$

we shall have $A_n \cdot \alpha_1 \cdots \alpha_n = A_n \cdot \alpha_1' \cdots \alpha_n'$.

The process would of course break down in case

$$A_0 \cdot \alpha_1 \alpha_1 = 0, \qquad A_0 \cdot A_2 \cdot \alpha_1 \alpha_2 A_2 \cdot \alpha_1 \alpha_2 = 0,$$

etc. But these are *singular vectors* and are excluded from the cases that are under consideration. This process of orthogonalizing vectors is of importance in the study of sets of functions, integral equations, etc. If the A_0 process is defined as in the General Analysis of Moore, so that it is a Hermitian product, the vectors may always be orthogonalized.

10. In case ρ is linearly expressible in terms of $\beta_1, \cdots, \beta_n$, since we have for any set of n vectors $\alpha_1, \cdots, \alpha_n$ (writing merely α for $\alpha_1 \alpha_2 \cdots \alpha_n$ and β for $\beta_1 \beta_2 \cdots \beta_n$)

$$\begin{aligned}
A_1 \cdot A_n \cdot \alpha A_{n+1} \cdot \beta \rho = 0 = (A_0 \cdot A_n \cdot \alpha A_n \beta)\rho \\
- (A_0 \cdot A_n \cdot \alpha A_n \cdot \beta_1 \cdots \beta_{n-1} \rho)\beta_n + \cdots
\end{aligned}$$

it follows that we have a means of expanding ρ in terms of the vectors $\beta_1, \beta_2, \cdots, \beta_n$ if $A_0 \cdot A_n \cdot \alpha A_n \cdot \beta \neq 0$; that is, if the set of α's has rank n with regard to the set of β's. Transposing we have

$$\begin{aligned}
\rho = \beta_n \frac{A_0 \cdot \alpha_n \cdot A_n \cdot \beta_1 \cdots \beta_{n-1} \rho}{A_0 \cdot A_n \cdot \alpha A_n \cdot \beta} \\
- \beta_{n-1} \frac{A_0 \cdot A_n \cdot \alpha A_n \cdot \beta_1 \cdots \beta_{n-2} \beta_n \rho}{A_0 \cdot A_n \cdot \alpha A_n \cdot \beta} + \cdots.
\end{aligned}$$

We now set

$$\beta^i = (-)^{i-1} A_1 \cdot A_{n-1} \cdot \beta_1 \cdots \beta_n A_n \cdot \alpha_1 \cdots \alpha_n / A_0 \cdot A_n \cdot \alpha A_n \cdot \beta;$$

β^i is called the *complement of* β_i *as to the vectors* α. In case the vectors α

coincided each to each with the corresponding vectors β, then the complements would be called the *adjuncts* of the vectors β, that is, the adjuncts are complements in the complex of the β's themselves. Evidently

$$A_0 \cdot \beta_i \beta^j = \delta_{ij},$$

where δ_{ij} is 1 if $i = j$, and 0 if $i \neq j$. The expression of ρ now becomes

$$\rho = \sum_i \beta_i A_0 \cdot \beta^i \rho = \sum_i \beta^i A_0 \cdot \beta_i \rho.$$

This is the generalized form for the expansion of ρ commonly used in quaternions.

The process used above for two complexes is called *biorthogonalization* of the vectors β as to the vectors α. The complexes need not be the same, but must have their rank with regard to each other equal to the order of the complex. That is to say, the vectors α must be expressible linearly in terms of the common incident complex and other vectors orthogonal to the incident complex. The same is true for the vectors β. The vectors outside of the incident complex, however, may be quite different. As an example we may show that the sets of vectors

$$\beta_1, \quad A_1 \cdot \alpha_1 A_2 \cdot \beta_1 \beta_2, \quad A_1 \cdot A_2 \cdot \alpha_1 \alpha_2 A_3 \cdot \beta_1 \beta_2 \beta_3, \quad A_1 \cdot A_3 \cdot \alpha_1 \alpha_2 \alpha_3 A_4 \cdot \beta_1 \beta_2 \beta_3 \beta_4, \quad \cdots,$$

$$\alpha_1, \quad A_1 \cdot \beta_1 A_2 \cdot \alpha_1 \alpha_2, \quad A_1 \cdot A_2 \cdot \beta_1 \beta_2 A_3 \cdot \alpha_1 \alpha_2 \alpha_3, \quad A_1 \cdot A_3 \cdot \beta_1 \beta_2 \beta_3 A_4 \cdot \alpha_1 \alpha_2 \alpha_3 \alpha_4, \quad \cdots$$

are biorthogonalized sets.

11. We may now give the complete definition that states the result of using either the accretive or the decretive process on two alternants* (counting $A_1 \cdot \xi$ as an alternant by definition). It could not have been stated before since some of the expressions on the right would have to be defined first. We define for all i, j and any vectors α, β,

$$A_{i+j-2k} \cdot A_i \cdot \alpha_1 \cdots \alpha_i A_j \cdot \beta_1 \cdots \beta_j = \sum_{i_1 \cdots i_k j_1 \cdots j_k} (-)^{\sum i_s - \sum j_t + k(i-k)} A_{i+j-2k} \cdot \alpha_1 \cdot \cdots \cdot \alpha_i \beta_1 \cdots \beta_j \times (A_0 \cdot A_k \cdot \alpha_{i_1} \cdots \alpha_{i_k} A \cdot \beta_{j_1} \cdots \beta_{j_k})$$

where † we have $k \leqq i, k \leqq j$, that is, $i + j - 2k \geqq |i - j|$. The sign may be determined directly without using the factor containing the sign, by counting the total number of inversions, provided however that in the scalar factor $A_0 \cdot ()()$ the vectors α precede the vectors β. This definition includes all

* No special name is needed, as the next division shows.

† This $\times$ sign merely means that the expression should be written on one line.

the others given, as particular cases. In case $k \leqq \frac{1}{2}i$, $i \leqq j$, and $\beta_{i_a} = \alpha_{i_a}$, $i_a = 1, 2, \cdots, i$, the expression vanishes. In case $k = 0$ we have but a single term, the sign being $+$. If $k = 0$ we have the accretive process. If the order of the two alternants is changed it is evident that the result is multiplied by $(-)^{ij+k}$.

The vector just defined, of grade $i + j - 2k$, is the alternant which defines the region or space of order $i + j - 2k$ containing all the regions of the alternants of order i and j determined by omitting common incident regions of order j.

4. The Hamilton product

12. **Hamilton product of two vectors.** Let V_n and V_m be any two vectors of grades n, m respectively, that is, in the complex of a system of alternants all of grade n and the complex of a system of alternants of grade m respectively. Then we define $V_{m+n-2k} \cdot V_m V_n = \sum A_{m+n-2k} \cdot A'_m A''_n x_m y_n$, where $V_m = \sum x_m A'_m$, and $V_n = \sum y_n A''_n$. This merely states the distributive law. We now define the Hamilton product (also distributive) for two vectors of grades m, n, indicating it by the sign * temporarily. We shall see later that no special mark is needed. We remember a scalar is not a vector, though written A_0.

$$V_m{}^* V_n = \sum_{k=0}^{k=|m-n|} V_{m+n-2k} \cdot V_m V_n, \quad 0 < m, 0 < n, \quad V_m{}^* x = x V_m = x^* V_m.$$

From this definition we have first, since $A_1 \cdot \alpha_1 = \alpha_1$,

$$\alpha_1{}^* \alpha_2 = A_0 \cdot \alpha_1 \alpha_2 + A_2 \cdot \alpha_1 \alpha_2.$$

Since the two parts of this product are of different grades, there is nothing to interfere with our using the same symbols in front of the Hamilton product as selective symbols to indicate the different parts. That is, we may use

$$\begin{aligned} A_0 \cdot \alpha_1{}^* \alpha_2 \quad &\text{and} \quad A_0 \cdot \alpha_1 \alpha_2, \\ A_2 \cdot \alpha_1{}^* \alpha_2 \quad &\text{and} \quad A_2 \cdot \alpha_1 \alpha_2 \end{aligned}$$

indifferently. We shall see that this is a procedure which is general.

Again we have

$$\begin{aligned} \alpha_1{}^* A_2 \cdot \alpha_2 \alpha_3 &= A_1 \cdot \alpha_1 A_2 \cdot \alpha_2 \alpha_3 + A_3 \cdot \alpha_1 A_2 \cdot \alpha_2 \alpha_3 \\ &= \alpha_3 A_0 \cdot \alpha_1 \alpha_2 - \alpha_2 A_0 \cdot \alpha_1 \alpha_3 + A_3 \cdot \alpha_1 \alpha_2 \alpha_3. \end{aligned}$$

Since $\alpha_1{}^* A_0 \cdot \alpha_2 \alpha_3 = \alpha_1 A_0 \cdot \alpha_2 \alpha_3$, by adding, we have

$$\alpha_1{}^* (\alpha_2{}^* \alpha_3) = \alpha_1 A_0 \cdot \alpha_2 \alpha_3 - \alpha_2 A_0 \cdot \alpha_1 \alpha_3 + \alpha_3 A_0 \cdot \alpha_1 \alpha_2 + A_3 \cdot \alpha_1 \alpha_2 \alpha_3.$$

This product consists then of two vectors, one of grade 1, the other of grade 3. If we use A_2 and A_3 in front of a Hamilton product as selective symbols, then we may define

$$\begin{aligned} A_1 \cdot \alpha_1{}^* (\alpha_2{}^* \alpha_3) &= \alpha_1 A_0 \cdot \alpha_2 \alpha_3 - \alpha_2 A_0 \cdot \alpha_1 \alpha_3 + \alpha_3 A_0 \cdot \alpha_1 \alpha_2, \\ A_3 \cdot \alpha_1{}^* (\alpha_2{}^* \alpha_3) &= A_3 \cdot \alpha_1 \alpha_2 \alpha_3. \end{aligned}$$

Evidently

$$A_1 \cdot \alpha_1 {}^*(\alpha_2 {}^* \alpha_3) = A_1 \cdot \alpha_3 {}^*(\alpha_2 {}^* \alpha_1).$$

We have

$$\alpha_1 {}^*(\alpha_2 {}^* \alpha_3) = A_1 \cdot \alpha_1 {}^*(\alpha_2 {}^* \alpha_3) + A_3 \cdot \alpha_1 \alpha_2 \alpha_3 = (\alpha_1 {}^* \alpha_2) {}^* \alpha_3.$$

The vector of grade 3 in this product is familiar. The vector of grade 1, however, is a new function of α_1, α_2, and α_3. We may indicate it as a function of α_1, α_2, α_3 by using the same symbol, omitting the multiplication sign*. The alternants we have already written in that manner, the mere juxtaposition of the vectors in $A_n \cdot \alpha_1 \cdots \alpha_n$ not implying up to this point anything in the nature of a product. It has meant so far nothing more than a certain defined function $A_n \cdot (\alpha_1, \alpha_2, \cdots, \alpha_n)$ of n vectors, of first grade. Hence we will now define a new system of functions recurrently:

$$(1) \qquad A_n \cdot \alpha_1 \alpha_2 \cdots \alpha_{n+2i} = A_n \cdot \alpha_1 A_{n-1} \cdot \alpha_2 \cdots \alpha_{n+2i} + A_n \cdot \alpha_1 A_{n+1} \cdot \alpha_2 \cdots \alpha_{n+2i}, \qquad 0 < n,$$

where $0 \leqq i$. (The last term vanishes when $i = 0$, since $A_{n+1} \cdot \alpha_2 \cdots \alpha_n = 0$, for $n + 1 > n$.)

$$(2) \qquad A_0 \cdot \alpha_1 \alpha_2 \cdots \alpha_{2m} = A_0 \cdot \alpha_1 A_1 \cdot \alpha_2 \cdots \alpha_{2m}.$$

For instance

$$A_1 \cdot \alpha_1 \alpha_2 \alpha_3 = A_1 \cdot \alpha_1 A_0 \cdot \alpha_2 \alpha_3 + A_1 \cdot \alpha_1 A_2 \cdot \alpha_2 \alpha_3.$$

In the case of functional space the number i indicates the number of integrations we perform, there being in all $n + 2i$ parameters which are set equal in i pairs in every possible manner, and then each pair integrated, leaving in every resulting term n parameters as indicated always by the subscript of the A. We see that now A has ceased to have only an alternating significance, retaining that meaning when the number of vectors following is the same as the subscript on the A. But when the number is greater (it never can be less), then A no longer is an alternating sign. The functions, however, do permit certain permutations depending upon the grade, as will be shown.

13. We shall now prove that the definition just given for A_n, when the number of vectors exceeds n, leads to an expansion in alternants of grade n. That is, we shall prove the theorem:

$$A_{n-2i} \cdot \alpha_1 \cdots \alpha_n = \sum_{j_1 j_2 \cdots j_{2i}} (-)^{\mathfrak{M}_{i+1}} A_0 \cdot \alpha_{j_1} \alpha_{j_2} \cdots \alpha_{j_{2i}} \times A_{n-2i} \cdot \alpha_1 \cdot \cdots \cdot \alpha_n \quad (a = 1, 2, \cdots, 2i).$$

As a matter of fact again the signs follow the determinant rule, provided we place the subscripts after A_0 in natural order, that is, have $j_1 \leqq j_2 \leqq \cdots \leqq j_{2i}$.

For let us suppose this formula holds for all values of n up to and including

$n = t$, and for all values of i, such that $2i \leqq t$. Then it follows that

$$A_1 \cdot \alpha_1 \alpha_2 \cdots \alpha_t = \alpha_1 A_0 \cdot \alpha_2 \cdots \alpha_t - \alpha_2 A_0 \cdot \alpha_1 \cdots \alpha_t + \cdots + \alpha_t A_0 \cdot \alpha_1 \cdots \alpha_{t-1}$$

whenever t is odd, and further when t is odd, so that $t + 1$ is even, from the definition in § 12,

$$\begin{aligned} A_0 \cdot \alpha_1 \alpha_2 \cdots \alpha_{t+1} &= A_0 \cdot \alpha_1 A_1 \cdot \alpha_2 \cdots \alpha_{t+1} \\ &= A_0 \cdot \alpha_1 \alpha_2 A_0 \cdot \alpha_3 \cdots \alpha_{t+1} - A_0 \cdot \alpha_1 \alpha_3 A_0 \cdot \alpha_2 \alpha_4 \cdots \alpha_{t+1} \cdots . \end{aligned}$$

Now by definition whether t is odd or even

$$\begin{aligned} A_{t+1-2j} \cdot \alpha_1 \alpha_2 \cdots \alpha_{t+1} &= A_{t+1-2j} \cdot \alpha_1 A_{t+2-2j} \cdot \alpha_2 \cdots \alpha_{t+1} + A_{t+1-2j} \cdot \alpha_1 A_{t-2j} \cdot \alpha_2 \cdots \alpha_{t+1}, \end{aligned}$$

where $2j \leqq t + 1$. But each term on the right contains a case that comes under the hypothesis since there are only t vectors concerned in the second A. Hence we may reduce the right-hand member, and

$$\begin{aligned} A_{t+1-2j} \cdot \alpha_1 \alpha_2 \cdots \alpha_{t+1} &= A_{t+1-2j} \cdot \alpha_1 \sum_{k_1 \cdots k_{2j-2}} A_{t+2-2j} \cdot \alpha_2 \cdot \cdots \cdot \alpha_{t+1} A_0 \cdot \alpha_{k_1} \cdots \alpha_{k_{2j-2}} (-)^{j-1+\Sigma k} \\ &\quad + A_{t+1-2j} \cdot \alpha_1 \sum_{h_1 \cdots h_{2j}} A_{t-2j} \cdot \alpha_2 \cdot \cdots \cdot \alpha_{t+1} A_0 \cdot \alpha_{h_1} \cdots \alpha_{h_{2j}} (-)^{j+\Sigma h} \end{aligned}$$

where $k_1 < k_2 < \cdots < k_{2j-2}$, $h_1 < h_2 < \cdots < h_{2j}$.

The first term reduces to

$$\sum_{k_0 k_1 \cdots k_{2j-2}} A_{t+1-2j} \cdot \alpha_2 \cdot \cdots \cdot \alpha_{t+1} A_0 \cdot \alpha_1 \alpha_{k_0} A_0 \cdot \alpha_{k_1} \cdots \alpha_{k_{2j-2}} (-)^{j-1+\Sigma k+k_0+r}$$

where r is the number of subscripts $k_1, k_2, \cdots, k_r < k_0$ in the term, and $k_0 \neq k_1, k_2, \cdots, k_{2j-2}$, $k_0 = 2, 3, \cdots, t + 1$ in different terms. This reduces to

$$\sum_{g, k_1 \cdots k_{2j-1}} A_{t+1-2j} \cdot \alpha_2 \alpha_3 \cdot \cdots \cdot \alpha_{t+1} A_0 \cdot \alpha_1 \alpha_2 \cdots \alpha_{k_{2j-1}} (-)^{j+1+g+\Sigma k}$$

in which the subscripts are now rearranged so that $g < k_1 < k_2 < \cdots < k_{2j-2}$. No $k = 1$. In these terms α_1 is missing with $2j - 1$ others, or $2j$ including α_1. The second term reduces to

$$\sum_{h_1 \cdots h_{2j}} A_{t+1-2j} \cdot \alpha_1 \alpha_2 \cdot \cdots \cdot \alpha_{t+1} A_0 \cdot \alpha_{h_1} \cdots \alpha_{h_{2j}} (-)^{j+\Sigma h}$$

where no h is 1, and $h_1 < h_2 < \cdots < h_{2j}$.

In these terms α_1 is present, and $2j$ excluding α_1 are missing. But these two sets of terms together give exactly

$$A_{t+1-2j} \cdot \alpha_1 \alpha_2 \cdots \alpha_{t+1} = \sum_{k_1 \cdots k_{2j}} A_{t+1-2j} \cdot \alpha_1 \cdot \cdots \cdot \alpha_{t+1} A_0 \cdot \alpha_{k_1} \cdots \alpha_{k_{2j}} (-)^{j+\Sigma k}$$

and hence the theorem holds for the next value of $n = t + 1$ and for all values of j that are possible. Since, however, we know the theorem is true for $n = 1$, $n = 2$, $n = 3$, it holds for all cases of n an integer. The signs follow the determinant rule provided that under $A_0 \cdot$ the letters are in their natural order, that is, as on the left. We see now from the definition of the Hamilton product of two vectors of any grades that

$$\begin{aligned}\alpha_1 {}^*[A_n \cdot \alpha_2 \cdots \alpha_{n+1} + A_{n-2} \cdot \alpha_2 \cdots \alpha_{n+1} \\ + \cdots + A_1 \quad \text{or} \quad A_0 \cdot \alpha_2 \cdots \alpha_{n+1}] \\ = A_{n+1} \cdot \alpha_1 \alpha_2 \cdots \alpha_{n+1} + A_{n-1} \cdot \alpha_1 A_n \cdot \alpha_2 \cdots \alpha_{n+1} \\ + \cdots + A_{n-1} \cdot \alpha_1 A_{n-2} \cdot \alpha_2 \cdots \alpha_{n+1} + \cdots \\ = A_{n+1} \cdot \alpha_1 \cdots \alpha_{n+1} + A_{n-1} \cdot \alpha_1 \cdots \alpha_{n+1} \\ + \cdots + A_0 \quad \text{or} \quad A_1 \cdot \alpha_1 \cdots \alpha_{n+1}.\end{aligned}$$

Now we know already that

$$\alpha_{n-1} {}^* \alpha_n = A_2 \cdot \alpha_{n-1} \alpha_n + A_0 \cdot \alpha_{n-1} \alpha_n$$

and

$$\alpha_{n-2} {}^* (\alpha_{n-1} {}^* \alpha_n) = A_3 \cdot \alpha_{n-2} \alpha_{n-1} \alpha_n + A_1 \cdot \alpha_{n-2} \alpha_{n-1} \alpha_n.$$

It is evident from the theorem just proved that we must have the general Hamilton product (in which the multiplier is set at the *left*) given by the theorem

$$\alpha_1 {}^* [\alpha_2 {}^* (\cdots {}^* \alpha_n)] = A_n \cdot \alpha_1 \cdots \alpha_n + A_{n-2} \cdot \alpha_1 \cdots \alpha_n + \cdots.$$

If the A's are inserted in front of this product they are understood as selective symbols that pick out the term on the right of corresponding grade. We have now to show that this product is *associative*, that is, we may group the vectors in the order in which they come, in any manner, and find that the Hamilton product of the groups is the Hamilton product of the whole set. For instance

$$[\alpha_1 {}^* (\alpha_2 {}^* \alpha_3)] {}^* (\alpha_4 {}^* \alpha_5) = \alpha_1 {}^* [(\alpha_2 {}^* \alpha_3) {}^* (\alpha_4 {}^* \alpha_5)] = \alpha_1 {}^* [\alpha_2 {}^* (\alpha_3 {}^* \overline{\alpha_4 {}^* \alpha_5})].$$

14. **Associativity of the Hamilton product.** We shall prove this by showing that whether we multiply successively α_n by α_{n-1}, then α_{n-2}, $\cdots$, by α_1, or whether we form the product of α_n by α_{n-1}, $\cdots$, by α_i, and then the product of this by the product of α_i multiplied successively by α_{i-1}, α_{i-2}, $\cdots$, by α_1, we have the same result. That is, we are to prove that

$$\begin{aligned}[A_i \cdot \alpha_1 \cdots \alpha_i + A_{i-2} \cdot \alpha_1 \cdots \alpha_i + \cdots] \\ {}^* [A_{n-i} \cdot \alpha_{i+1} \cdots \alpha_n + A_{n-i-2} \cdot \alpha_{i+1} \cdots \alpha_n] \\ = A_n \cdot \alpha_1 \cdots \alpha_n + A_{n-2} \cdot \alpha_1 \cdots \alpha_n + \cdots.\end{aligned}$$

In the first place let $n = i + j$. Then

$$A_{i-2q} \cdot \alpha_1 \cdots \alpha_i = \sum (-)^{\Sigma i_q + q} A_{i-2q} (i_1, \cdots, i_{2q}) A_0 \cdot \alpha_{i_1} \cdots \alpha_{i_{2q}},$$

where the subscripts written inside the parenthesis will indicate the *missing* vectors from $\alpha_1, \cdots, \alpha_i$. Also we have

$$A_{j-2m}\,\alpha_{i+1}\cdots\alpha_{i+j} = \sum(-)^{\Sigma j_i+m} A_{j-2m}(j_1,\cdots,j_{2m})\,A_0\cdot\alpha_{j_1}\cdots\alpha_{j_{2m}},$$

$$\begin{aligned} A_{i+j-2(m+p+q)}\cdot A_{i-2q}\cdot\alpha_1\cdots\alpha_i\,A_{j-2m}\cdot\alpha_{i+1}\cdots\alpha_{i+j} \\ = \sum(-)^{\Sigma i_s+\Sigma j_i+q+m} A_{i+j-2(m+p+q)}\,A_{i-2q}(i_1,\cdots,i_{2q}) \\ \times A_{j-2m}(j_1,\cdots,j_{2m})\,A_0\cdot\alpha_{i_1}\cdots\alpha_{i_{2q}}A_0\cdot\alpha_{j_1}\cdots\alpha_{j_{2m}} \\ = \sum(-)^{\Sigma i_s+\Sigma j_i+q+m+p} A_{i+j-2(m+p+q)}\,\alpha_1\,\alpha_2\cdot\cdots\cdots\cdots\cdot\alpha_{i+j} \\ \times A_0\cdot\alpha_{i_1}\cdots\alpha_{i_{2q}}A_0\cdot\alpha_{j_1}\cdots\alpha_{j_{2m}}A_0 \\ \cdot A_p\,\alpha_{i_{2q+1}}\cdots\alpha_{i_{2q+p}}A_p\cdot\alpha_{j_{2m+1}}\cdots\alpha_{j_{2m+p}}. \end{aligned}$$

Let $r = m + p + q$; then we may condense the expression just written into

$$\begin{aligned} \sum(-)^r A_{i+j-2r}(i_1\cdots i_{2q+p}\,j_1\cdots j_{2m+p})\,A_0\cdot A_p\cdot\alpha_{i_1} \\ \cdots\alpha_{i_{2q+p}}A_p\cdot\alpha_{j_1}\cdots\alpha_{j_{2m+p}}. \end{aligned}$$

If now we set $q + m = h$, $p = r - h$, and holding h fixed let q and m take all possible values, we have from the above

$$\begin{aligned} \sum_{h=0}^{h=r}\sum_{t=0,\,m=h}^{t=h,\,m=0} A_{i+j-2r}\cdot A_{i-2t}\cdot\alpha_1\cdots\alpha_i\,A_{j-2m}\,\alpha_{i+1}\cdots\alpha_{i+j} \\ = \sum_h(-)^r A_{i+j-2r}\,\alpha_1\cdot\cdots\cdot\alpha_i\,\alpha_{i+1}\cdot\cdots\cdot\alpha_j\,A_0\cdot A_{r-h}\cdot\alpha_{i_1} \\ \cdots\alpha_{i_s}A_{r-h}\cdot\alpha_{j_1}\cdots\alpha_{j_t}, \end{aligned}$$

where we must have $s + t = 2r$ and $r - h \leqq s \leqq r + h$. But finally summing as to h, we have exactly the sum

$$\sum_r A_{i+j-2r}\,\alpha_1\cdots\alpha_i\,\alpha_{i+1}\cdots\alpha_{i+j},$$

which is the expansion of the Hamiltonian product of $\alpha_1, \alpha_2, \cdots, \alpha_{i+j}$ in the order written. On the left the final sums give the Hamiltonian products of $\alpha_1, \alpha_2, \cdots, \alpha_i$ and $\alpha_{i+1}, \cdots, \alpha_{i+j}$.

We see now that the special use of a multiplication sign is unnecessary and that in all the functions we have built up, with the use of the A and subscripts, the array of vectors may be considered to be a Hamilton product, and the A a selective sign for the vector of the indicated grade. The remarkable simplicity of this result enables us to dispense with any very large collection of formulas for transforming expressions in this general vector calculus. We may break up by associativity any expression under an A and arrive at various reductions in this way. As instances

$$\begin{aligned} A_0\cdot\alpha_1\,\alpha_2\,\alpha_3\,\alpha_4\,\alpha_5\,\alpha_6 &= A_0\cdot(\alpha_1\,\alpha_2\,\alpha_3\,\alpha_4)(\alpha_5\,\alpha_6) \\ &= A_0\cdot[(A_4+A_2+A_0)\cdot\alpha_1\,\alpha_2\,\alpha_3\,\alpha_4][A_2+A_0]\cdot\alpha_5\,\alpha_6 \\ &= A_0\cdot A_2\cdot\alpha_1\,\alpha_2\,\alpha_3\,\alpha_4\,A_2\cdot\alpha_5\,\alpha_6 + A_0\cdot\alpha_1\,\alpha_2\,\alpha_3\,\alpha_4\,A_0\cdot\alpha_5\,\alpha_6; \end{aligned}$$

$$\begin{aligned} A_2\cdot(A_2\cdot\alpha_1\,\alpha_2\,A_4\cdot\beta_1\beta_2\,\beta_3\,\beta_4) \\ = A_2\,(\alpha_1\,\alpha_2 - A_0\cdot\alpha_1\,\alpha_2)\,A_4\cdot\beta_1\beta_2\,\beta_3\beta_4 = A_2\cdot(\alpha_1\,\alpha_2\,A_4\cdot\beta_1\beta_2\,\beta_3\,\beta_4); \end{aligned}$$

$$\begin{aligned} A_0\cdot\alpha_1\,\alpha_2\,\alpha_3\,\alpha_4 = A_0\cdot\alpha_1\,\alpha_2\,A_0\cdot\alpha_3\,\alpha_4 \\ - A_0\cdot\alpha_1\,\alpha_3\,A_0\cdot\alpha_2\,\alpha_4 + A_0\cdot\alpha_1\,\alpha_4\,A_0\cdot\alpha_2\,\alpha_3. \end{aligned}$$

Expansions of this sort are unlimited in number.* Other useful formulas are easily derived by combinations, as

$$A_n \cdot \alpha A_{n-1} \beta A_n \cdot \gamma_1 \cdots \gamma_n + A_n \cdot \beta A_{n+1} \cdot \alpha A_n \cdot \gamma_1 \cdots \gamma_n = A_n \cdot \alpha\beta A_n \cdot \gamma_1 \cdots \gamma_n;$$

$$A_1 \cdot A_n \cdot \alpha_1 \cdots \alpha_n A_{n+1} \rho\beta_1 \cdots \beta_n = (-)^n \rho A_0 \cdot A_n \cdot \alpha_1 \cdots \alpha_n A_n \cdot \beta_1 \cdots \beta_n - \sum_j \beta_j A_0 \cdot A_n \cdot \alpha_1 \cdots \alpha_n A_n \cdot \rho\beta_1 \cdots \beta_n;$$

$$A_1 \cdot \alpha_1 A_2 \cdot \alpha_2 \alpha_3 + A_1 \cdot \alpha_2 A_2 \cdot \alpha_3 \alpha_1 + A_1 \cdot \alpha_3 A_2 \cdot \alpha_1 \alpha_2 = 0;$$

$$A_2 \cdot \alpha_1 A_3 \cdot \alpha_2 \alpha_3 \alpha_4 - A_2 \cdot \alpha_2 A_3 \cdot \alpha_1 \alpha_3 \alpha_4 + A_2 \cdot \alpha_3 A_3 \cdot \alpha_1 \alpha_2 \alpha_4 - A_2 \cdot \alpha_4 A_3 \cdot \alpha_1 \alpha_2 \alpha_3 = 0.$$

15. There are other ways by which we might have developed the subject. We might have taken a method similar to that of Clifford and defined the symbols α, β, etc., which we have called vectors, as the hypernumbers defined by the ratios of geometric vectors; but this method (and others) has great disadvantages in the end. We might of course have defined a general or indeterminate product and derived from this expressions which would have been our A expressions. Some of these will be exhibited later. We should have difficulties also this way. From a long consideration of the problem the method of development used was settled on as the most desirable from a good many points of view.

16. We proceed to develop a few theorems useful in applications. If we consider two complexes of order n each, defined by $\alpha_1, \cdots, \alpha_n$ and $\beta_1, \cdots, \beta_n$, and a vector ρ, the part of ρ in the incident complex of the two (say ρ') is given by

$$\rho' A_0 \cdot A_n \cdot \alpha A_n \cdot \beta = A_1 \cdot (A_{n-1} \cdot \rho A_n \alpha) A_n \cdot \beta.$$

For it is easy to see that

$$A_{n-1} \cdot \rho A_n \cdot \alpha_1 \cdots \alpha_n = \sum_i A_0 \cdot \alpha_i \rho A_{n-1} \cdot \alpha_1 \cdots \alpha_n (-)^{i-1},$$

and therefore the right side gives $\sum_i A_0 \cdot \alpha_i \rho \cdot \alpha^i \cdot A_0 \cdot A_n \cdot \alpha A_n \cdot \beta$, when we set for the complement of α_i

$$\alpha^i = (-)^{i-1} \cdot A_1 \cdot A_{n-1} \cdot \alpha_1 \cdots \alpha_n A_n \cdot \beta_1 \cdots \beta_n / A_0 \cdot A_n \cdot \alpha A_n \cdot \beta.$$

But the form we now have is that part of ρ which lies in the complex of $\alpha^1, \alpha^2, \cdots, \alpha^n$, that is, in the incident complex. It follows that

$$\rho - \frac{A_1 \cdot (A_{n-1} \cdot \rho A_n \cdot \alpha) A_n \cdot \beta}{A_0 \cdot A_n \cdot \alpha A_n \cdot \beta}$$

*Conversely every expression $A_m \cdot \alpha_1 \cdots \alpha_n$ may be expressed as a sum of products alone without any A's. See page 220.

is the part of ρ (its *projection*) *exterior* to the incident space. It may be also written in the form

$$\frac{A_1 \cdot [\rho A_0 \cdot A_n \cdot \alpha A_n \cdot \beta - (A_{n-1} \cdot \rho A_n \cdot \alpha) A_n \cdot \beta]}{A_0 \cdot A_n \cdot \alpha A_n \cdot \beta}.$$

17. Another useful form is the following. Since $A_1 \cdot A_{n-1} \cdot \beta_1 \cdots \beta_{n-1} A_n \cdot \alpha$ is linear in $\alpha_1, \cdots, \alpha_n$, we see that if we consider a set of vectors defining a complex of order r, $\alpha_1, \cdots, \alpha_n, \alpha_{n+1}, \cdots, \alpha_r$, $n \leqq r$, then

$$A_{r+1} \cdot [A_1 \cdot A_{n-1} \cdot \beta_1 \cdots \beta_{n-1} A_n \cdot \alpha_1 \cdots \alpha_n] A_r \cdot \alpha_1 \cdots \alpha_r = 0$$

identically. This is true whatever the vectors β. Hence for any vectors $\gamma_1, \cdots, \gamma_{r+1}$

$$A_0 \cdot [A_{r+1} \cdot \gamma_1 \cdots \gamma_{r+1} A_{r+1} \cdot A_1 (A_{n-1} \cdot \beta_1 \cdots \beta_{n-1} A_n \cdot \alpha_1 \cdots \alpha_n) A_r \cdot \alpha_1 \cdots \alpha_r] = 0$$

identically. We may now change the order and have still identically

$$A_0 \cdot [A_{r+1} \cdot (A_r \cdot \alpha_1 \cdots \alpha_r A_1 \cdot A_{n-1} \cdot \beta_1 \cdots \beta_{n-1} A_n \cdot \alpha_1 \cdots \alpha_n) A_{r+1} \cdot \gamma_1 \cdots \gamma_{r+1}] = 0.$$

We may however drop the first A_{r+1} since A_0 in front prevents any additional terms from entering, and then for the same reason we may insert A_r as shown, arriving at a formula from which many useful formulas may be deduced: *

$$A_0 \cdot A_r \cdot \alpha_1 \cdots \alpha_r A_r \cdot [A_1 \cdot (A_{n-1} \cdot \beta_1 \cdots \beta_{n-1} A_n \cdot \alpha_1 \cdots \alpha_n) A_{r+1} \cdot \gamma_1 \cdots \gamma_{r+1}] = 0.$$

18. A similar formula may be deduced from

$$A_x \cdot A_y \cdot \beta_1 \cdots \beta_y A_{x+y} \cdot \alpha_1 \cdots \alpha_{x+y},$$

which is linear in $A_x \cdot \alpha_{i_1} \cdots \alpha_{i_x}$. If then we consider the set $\alpha_1, \cdots, \alpha_{x+y+z}$ where z may be zero, we have the identity

$$A_{2x+y+z} \cdot (A_x \cdot A_y \cdot \beta_1 \cdots \beta_y A_{x+y} \cdot \alpha_1 \cdots \alpha_{x+y}) A_{x+y+z} \cdot \alpha_1 \cdots \alpha_{x+y+z} = 0.$$

Hence whatever the vectors $\gamma_1, \cdots, \gamma_{2x+y+z}$, we have identically

$$A_0 \cdot [A_{2x+y+z} \cdot \gamma_1 \cdots \gamma_{2x+y+z} A_{2x+y+z} \cdot (A_x \cdot A_y \cdot \beta_1 \cdots \beta_y A_{x+y} \cdot \alpha_1 \cdots \alpha_{x+y}) A_{x+y+z} \cdot \alpha_1 \cdots \alpha_{x+y+z}] = 0,$$

whence for all β's and γ's we have

$$A_0 \cdot [A_{x+y+z} \cdot \alpha_1 \cdots \alpha_{x+y+z} A_x \cdot (A_y \cdot \beta_1 \cdots \beta_y A_{x+y} \cdot \alpha_1 \cdots \alpha_{x+y})] A_{2x+y+z} \cdot \gamma_1 \cdots \gamma_{2x+y+z} = 0.$$

Particular cases occur for $z = 0$.

* Volterra, Rice Institute Pamphlets, vol. 4 (1917), No. 1, p. 58 (3).

19. In the expression

$$A_{2r+s-2x} \cdot A_r \cdot \alpha_1 \cdots \alpha_r \, A_{r+s} \cdot \alpha_1 \cdots \alpha_r \, \beta_1 \cdots \beta_s$$
$$= \sum \pm A_0 \cdot A_x \cdot \alpha_{i_1} \cdots \alpha_{i_x} A_x \cdot \alpha_{j_1} \cdots \alpha_{j_y} \beta_{k_1}$$
$$\cdots \beta_{k_{x-y}} A_{2r+s-2x} \cdot \alpha_1 \underset{i_1 \cdots i_x}{\cdots} \alpha_r \, \alpha_1 \underset{j_1 \cdots j_y}{\cdots} \alpha_r \, \beta_1 \underset{k_1 \cdots k_{x-y}}{\cdots} \beta_s$$

the vectors α that enter the right-hand side are all different, and there are not more than r, hence we have $2r - x - y \leqq r$ or $r \leqq x + y$. Further if the j set and the i set do not together make up the whole r set, then there are inevitably duplicates in the last factor which would cause it to vanish. Hence the i set and the j set must cover all the r vectors $\alpha_1, \cdots, \alpha_r$, though they may overlap, in which case in the last factor the overlap would be missing. In each A we may reduce with no alteration in value by substituting an orthogonal set for the vectors appearing there (indicated by accents). The factor A_0 would then vanish if it contained any vectors β. Hence the vectors of the i set and those of the j set must define the same complex. They must therefore each be the whole set $\alpha_1, \cdots, \alpha_r$. Hence for any value except $x = y = r$ the expression vanishes, and when $x = y = r$ it reduces to

$$A_s \cdot \beta'_1 \cdots \beta'_s .$$

20. If we consider the bordered determinant

$$\begin{vmatrix} A_0 \cdot \alpha\delta & A_0 \cdot \alpha\beta_1 & A_0 \cdot \alpha\beta_2 & \cdots & A_0 \cdot \alpha\beta_n \\ A_0 \cdot \gamma_1\delta & A_0 \cdot \gamma_1\beta_1 & A_0 \cdot \gamma_1\beta_2 & \cdots & A_0 \cdot \gamma_1\beta_n \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ A_0 \cdot \gamma_n\delta & A_0 \cdot \gamma_n\beta_1 & A_0 \cdot \gamma_n\beta_2 & \cdots & A_0 \cdot \gamma_n\beta_n \end{vmatrix}$$

we see that it may be condensed into

$$(-)^{\frac{1}{2}n(n+1)} A_0 \cdot A_{n+1} \cdot \delta\beta_1\beta_2 \cdots \beta_n \, A_{n+1} \cdot \alpha\gamma_1\gamma_2 \cdots \gamma_n .$$

If we expand by Cauchy's theorem the determinant is equal to

$$\Delta \cdot A_0 \cdot \alpha\delta - \sum A_0 \cdot \alpha\beta_i \, A_0 \cdot \delta\gamma_j \cdot \Delta_{ij} ,$$

where Δ is the determinant of β and γ alone, that is, the one bordered, and Δ_{ij} is the minor of Δ produced by omitting the column containing β_i and row containing γ_j. Now interchange α and δ and subtract. The first terms cancel and the others become

$$\sum A_0 \cdot A_2 \cdot \alpha\delta A_2 \cdot \beta_i \gamma_j \cdot \Delta_{ij} = \sum A_0 \cdot A_2 \cdot \alpha\delta A_2$$
$$\cdot \beta_i \gamma_j \, A_0 \cdot A_{n-1} \cdot \beta_1 \underset{\beta_i}{\cdots} \beta_n \, A_{n-1} \cdot \gamma_1 \underset{\gamma_j}{\cdots} \gamma_n \, (-)^{\frac{1}{2}(n-1)(n-2)+i+j} .$$

This is easily reduced to

$$(-)^{\frac{1}{2}(n-1)n+1} A_0 \cdot A_2 \cdot \alpha\delta A_2 \cdot A_n \cdot \beta_1 \cdots \beta_n \, A_n \cdot \gamma_1 \cdots \gamma_n .$$

Hence we have the useful theorem

$$A_0 \cdot A_{n+1} \cdot \alpha\beta_1 \cdots \beta_n A_{n+1} \cdot \delta\gamma_1 \cdots \gamma_n - A_0 \cdot A_{n+1} \cdot \delta\beta_1 \cdots \beta_n A_{n+1} \cdot \alpha\gamma_1 \cdots \gamma_n = (-)^n A_0 \cdot A_2 \cdot \alpha\delta A_2 \cdot A_n \cdot \beta A_n \cdot \gamma .$$

21. We notice that

$$A_n \cdot \beta_1 \cdots \beta_n{}^* \alpha = A_{n+1} \cdot (A_n \cdot \beta_1 \cdots \beta_n{}^* \alpha) + A_{n-1} \cdot (A_n \cdot \beta_1 \cdots \beta_n{}^* \alpha) = (-)^n [A_{n+1} \cdot \alpha\beta_1 \cdots \beta_n - A_{n-1} \cdot \alpha A_n \cdot \beta_1 \cdots \beta_n].$$

Hence we have since $A_{n+2} \cdot \alpha A_{n+1} \alpha() = 0$, and $A_{n-2} \cdot \alpha A_{n-1} \cdot \alpha() = 0$,

$$\begin{aligned} \alpha^* A_n \cdot \beta_1 \cdots \beta_n{}^* \alpha &= (-)^n [A_n \cdot \alpha A_{n+1} \cdot \alpha\beta_1 \cdots \beta_n - A_n \cdot \alpha A_{n-1} \cdot \alpha A_n \cdot \beta_1 \cdots \beta_n] \\ &= (-)^n [A_0 \cdot \alpha\alpha A_n \cdot \beta_1 \cdots \beta_n - 2A_0 \cdot \alpha\beta_1 \cdot A_n \cdot \alpha\beta_2 \cdots \beta_n + 2A_0 \cdot \alpha\beta_2 A_n \cdot \alpha\beta_1 \beta_3 \cdots \beta_n] \\ &= A_n \cdot \alpha\beta_1 \cdots \beta_n \alpha . \end{aligned}$$

It is apparent then that if we multiply an alternant of grade n by the same linear vector on the right and left, we have a vector of grade n, which is the part of grade n of the product of the original product by this vector on the right and the left.

Since we have $A_0 \cdot \alpha\beta_1 \cdots \beta_n \alpha = A_0 \cdot \alpha\alpha A_0 \cdot \beta_1 \cdots \beta_n$ and since

$$\alpha\alpha = A_0 \cdot \alpha\alpha, \quad \text{and} \quad \alpha \cdot A_0 \cdot \beta_1 \cdots \beta_n \cdot \alpha = \alpha\alpha A_0 \cdot \beta_1 \cdots \beta_n ,$$

it is evident that for any expression

$$A_m \cdot \beta_1 \cdots \beta_n \qquad (m \leqq n)$$

we have

$$\alpha^* A_m \cdot \beta_1 \cdots \beta_n{}^* \alpha = A_m \cdot \alpha\beta_1 \cdots \beta_n \alpha .$$

A vector of grade 1 which comes at both extremes of any vector of any grade of a product may be set outside the vector symbol of that grade. This enables us to make reductions. For instance

$$\begin{aligned} \alpha\beta\gamma\alpha &= \alpha A_2 \cdot \beta\gamma \cdot \alpha + \alpha\alpha A_0 \cdot \beta\gamma ; \\ \alpha\beta\gamma\alpha\delta &= \alpha\alpha A_0 \cdot \beta\gamma \cdot \delta + \alpha A_2 \cdot \beta\gamma \cdot \alpha\delta . \end{aligned}$$

This theorem is useful in studying invariancy under rotation about linear vectors. The papers of Joly and M'Aulay should be consulted.

It is evident that continued application of the theorem is possible. For instance

$$(\alpha_1 \cdots \alpha_n) V_m \cdot \beta_1 \cdots \beta_r (\alpha_n \cdots \alpha_1) = V_m \cdot \alpha_1 \cdots \alpha_n \beta_1 \cdots \beta_r \alpha_n \cdots \alpha_1 .$$

22. Since the formulas we are to consider in this section are homogeneous in all vectors on both sides, we will consider that the vectors are *unit vectors*, that is to say for each vector we have

$$\xi\xi = A_0 \cdot \xi\xi = 1 .$$

We may introduce multipliers easily to produce vectors not unit vectors. We will occasionally indicate the opposite of a vector by a stroke over the symbol, thus

$$-\alpha = \bar{\alpha}.$$

We have then $\alpha\bar{\alpha} = -1$. We shall use the stroke extended to cover more than one vector to indicate that the order is to be reversed and all the signs, thus

$$\overline{\alpha_1 \cdots \alpha_n} = \bar{\alpha}_n \cdots \bar{\alpha}_1.$$

This is called, in quaternions, *taking the conjugate*.

This operation will not affect the form $A_0 \cdot \alpha_1 \cdots \alpha_n$, for this vanishes when n is odd, and if $n = 2m$ we have by definition of $A_0 \cdot \alpha_1 \cdots \alpha_{2m}$

$$A_0 \cdot \alpha_1 \cdots \alpha_{2m} = \sum (-)^p A_0 \cdot \alpha_{i_1} \alpha_{j_1} A_0 \cdot \alpha_{i_2} \alpha_{j_2} \cdots A_0 \cdot \alpha_{i_m} \alpha_{j_m},$$

where the subscripts are so arranged that $i_x < j_x$ $(x = 1, 2, \cdots, m)$ and p is the number of inversions in the total collection of subscripts so arranged. It simplifies the count a little to arrange either the subscripts i or j in their natural order. It is evident now that as each factor may have its vectors interchanged, and as a change of both signs does not affect the value, we will have in any case

$$A_0 \cdot \overline{\alpha_1 \cdots \alpha_n} = A_0 \cdot \alpha_1 \cdots \alpha_n.$$

Hence if we remember the expansion of the form $A_m \cdot \alpha_1 \cdots \alpha_n$, we see that as every term is an alternant of grade m, and as

$$A_m \cdot \overline{\beta_1 \cdots \beta_m} = (-)^m A_m \cdot \beta_m \cdots \beta_1 = (-)^{\frac{1}{2}m(m+1)} A_m \cdot \beta_1 \cdots \beta_m,$$

we shall have

$$A_m \cdot \overline{\alpha_1 \cdots \alpha_n} = (-)^{\frac{1}{2}m(m+1)} A_m \cdot \alpha_1 \cdots \alpha_n.$$

Hence we have the four forms for any product $\pi = \alpha_1 \cdots \alpha_n$, according to form of n,

$$\begin{aligned} A_{4m} \cdot \bar{\pi} &= A_{4m} \cdot \pi; & -A_{4m+1} \cdot \bar{\pi} &= A_{4m+1} \cdot \pi; \\ -A_{4m+2} \cdot \bar{\pi} &= A_{4m+2} \cdot \pi; & A_{4m+3} \cdot \bar{\pi} &= A_{4m+3} \cdot \pi. \end{aligned}$$

It follows that

$$\bar{\pi} = (A_0 \cdot - A_1 \cdot - A_2 \cdot + A_3 \cdot + A_4 \cdot - A_5 \cdot - \cdots) \cdot \pi.$$

Hence

$$\begin{aligned} \tfrac{1}{2}(\pi + \bar{\pi}) &= (A_0 \cdot + A_4 \cdot + \cdots)\pi && \text{if } \pi \text{ is of even grade,} \\ \tfrac{1}{2}(\pi - \bar{\pi}) &= (A_2 \cdot + A_6 \cdot + \cdots)\pi && \text{if } \pi \text{ is of even grade,} \\ \tfrac{1}{2}(\pi + \bar{\pi}) &= (A_1 \cdot + A_5 \cdot + \cdots)\pi && \text{if } \pi \text{ is of odd grade,} \\ \tfrac{1}{2}(\pi - \bar{\pi}) &= (A_3 \cdot + A_7 \cdot + \cdots)\pi && \text{if } \pi \text{ is of odd grade.} \end{aligned}$$

We may also state the theorem thus:* in $A_0, A_4, \cdots$ we may reverse the order of the product, in $A_1, A_5, \cdots$ we may reverse the order, in $A_2, A_6, \cdots$ and in $A_3, A_7, \cdots$ we must change the sign if we reverse the order. We may also state it thus: the substitution $(1,n)(2,n-1)(3,n-2)\cdots$ leaves invariant $A_0, A_4, \cdots, A_1, A_5, \cdots$ and changes the sign of $A_2, A_6, \cdots, A_3, A_7, \cdots$.

Again we notice that if μ is a vector of grade m, then

$$\mu\alpha = A_{m+1}\cdot\mu\alpha + A_{m-1}\cdot\mu\alpha, \qquad \alpha\mu = A_{m+1}\cdot\alpha\mu + A_{m-1}\cdot\alpha\mu,$$

since all other grades vanish, and by the formula in § 11 we have when we interchange the vectors

$$\alpha\mu = (-)^m[A_{m+1}\cdot\mu\alpha - A_{m-1}\cdot\mu\alpha],$$
$$\mu\alpha = (-)^m[A_{m+1}\cdot\alpha\mu - A_{m-1}\cdot\alpha\mu].$$

This gives us at once

$$\mu\alpha + (-)^m\alpha\mu = 2A_{m+1}\cdot\mu\alpha = (-)^m 2A_{m+1}\cdot\alpha\mu,$$
$$\mu\alpha - (-)^m\alpha\mu = 2A_{m-1}\cdot\mu\alpha = (-)^{m-1}2A_{m-1}\cdot\alpha\mu.$$

We have from this, reverting to $\pi = \alpha_1\cdots\alpha_n$, and writing $\pi_s = A_s\cdot\pi$ for brevity,

$$\pi\alpha = \sum\pi_s\alpha = \sum(-)^s[A_{s+1}\cdot\alpha\pi_s - A_{s-1}\cdot\alpha\pi_s].$$

We also have

$$\alpha\pi = \sum\alpha\pi_s = \sum[A_{s+1}\alpha\pi_s + A_{s-1}\cdot\alpha\pi_s].$$

It is clearer if we write these out in full, retaining only the terms actually present,

$$\alpha\pi = A_0\cdot\alpha\pi_1 + A_1\cdot\alpha(\pi_0+\pi_2) + A_2\cdot\alpha(\pi_1+\pi_3) + A_3\cdot\alpha(\pi_2+\pi_4) + \cdots,$$
$$\pi\alpha = A_0\cdot\alpha\pi_1 + A_1\cdot\alpha(\pi_0-\pi_2) - A_2\cdot\alpha(\pi_1-\pi_3) + A_3\cdot\alpha(\pi_2-\pi_4) + \cdots.$$

Of course when π is even all the even terms are zero in these, and when π is odd all the odd terms are zero. We have from these, by combining,

$$\tfrac{1}{2}[\alpha\pi+\pi\alpha] = A_0\cdot\alpha\pi_1 + A_1\cdot\alpha\pi_0 + A_2\cdot\alpha\pi_3 + A_3\cdot\alpha\pi_2 + \cdots,$$
$$\tfrac{1}{2}[\alpha\pi-\pi\alpha] = A_1\cdot\alpha\pi_2 + A_2\cdot\alpha\pi_1 + A_3\cdot\alpha\pi_4 + A_4\cdot\alpha\pi_3 + \cdots.$$

We have a number of interesting results from these. For instance, taking A_0 of both sides we find that

$$A_0\cdot\pi\alpha = A_0\cdot\alpha\pi.$$

Hence A_0 is invariant under the substitution $(1,2,\cdots,n)$. This substitution taken with the other which leaves A_0 invariant gives a group of order $2n$.

* Since $A_0(\pi-\overline{\pi}) = 0$, $A_4\cdot(\pi-\overline{\pi}) = 0$, etc.

Since in the expanded form, where $n = 2m$, we see that there are $2^m(m!)$ substitutions which leave the individual terms unchanged, and there are $(2m)!$ substitutions altogether, there will be in the development of $A_0 \cdot \alpha_1 \cdots \alpha_n$

$$\frac{(2m)!}{2^m(m!)} = (2m-1)(2m-3)\cdots 3 \cdot 1 \quad \text{terms.}$$

There will evidently be for $n = 2m$ vectors of grade 1, $\frac{1}{2}(n-1)!$ different expressions $A_0 \cdot \alpha_1 \cdots \alpha_n$.

Again we have

$$\begin{aligned}
&A_1 \cdot \pi\alpha = A_1 \cdot \alpha\pi - 2A_1 \cdot \alpha\pi_2 = -A_1 \cdot \alpha\pi + 2\alpha\pi_0;\\
&A_2 \cdot \pi\alpha = A_2 \cdot \alpha\pi - 2A_2 \cdot \alpha\pi_1 = -A_2 \cdot \alpha\pi + 2A_2 \cdot \alpha\pi_3;\\
&A_3 \cdot \pi\alpha = A_3 \cdot \alpha\pi - 2A_3 \cdot \alpha\pi_4 = -A_3 \cdot \alpha\pi + 2A_3 \cdot \alpha\pi_2;\\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\
&A_m \cdot \pi\alpha = A_m \cdot \alpha\pi - 2A_m \cdot \alpha\pi_{m\pm1} = -A_m \cdot \alpha\pi + 2A_m \cdot \alpha\pi_{m\mp1},
\end{aligned}$$

where the upper signs are used if *m is odd*, the lower if *m is even*. Returning to our previous procedure we have

$$\begin{aligned}
\mu\alpha\beta = \mu A_0 \cdot \alpha\beta + \mu A_2 \cdot \alpha\beta = \mu A_0 \cdot \alpha\beta + A_{m-2} \cdot \mu A_2 \cdot \alpha\beta\\
+ A_m \cdot \mu A_2 \cdot \alpha\beta + A_{m+2} \cdot \mu A_2 \cdot \alpha\beta,\\
= \mu A_0 \cdot \alpha\beta + A_{m-2} \cdot (A_2 \cdot \alpha\beta)\,\mu\\
- A_m \cdot (A_2 \cdot \alpha\beta)\,\mu + A_{m+2} \cdot (A_2 \cdot \alpha\beta)\,\mu,
\end{aligned}$$

$$\alpha\beta\mu = \mu A_0 \cdot \alpha\beta + A_{m-2} \cdot (A_2 \cdot \alpha\beta)\,\mu + A_m \cdot (A_2 \cdot \alpha\beta)\,\mu + A_{m+2} \cdot (A_2 \cdot \alpha\beta)\,\mu.$$

We have at once

$$\mu\alpha\beta = \alpha\beta\mu - 2A_m(A_2 \cdot \alpha\beta)\,\mu,$$

whence easily

$$\pi\alpha\beta = \alpha\beta\pi - 2\sum A_m(A_2 \cdot \alpha\beta)\,\pi_m.$$

We need not proceed in this fashion, for it is simpler to make use of the conjugates, thus: let

$$\pi = \alpha_1 \cdots \alpha_n \qquad \text{and} \qquad \tau = \beta_1 \cdots \beta_p;$$

then

$$\pi\tau = \overline{\overline{\tau}\,\overline{\pi}} = \sum (-)^{\frac{1}{2}m(m+1)} A_m \cdot \overline{\tau}\,\overline{\pi},$$

and since

$$\tau\pi = \sum A_m \cdot \tau\pi,$$

we have the general formula

$$\pi\tau \pm \tau\pi = \sum A_m \cdot (\tau\pi \pm (-)^{\frac{1}{2}m(m+1)}\,\overline{\tau}\,\overline{\pi}).$$

Now by expanding each expression in the form

$$\pi = \sum \pi_s, \qquad \overline{\pi} = \sum (-)^{\frac{1}{2}s(s+1)}\,\pi_s,$$

and similarly for τ, and cancelling all terms that must give 0, we arrive at any of these forms directly. The terms retained are of the form $\tau_p \pi_s$ where

$$m \leqq |p - s|, \qquad p + s \leqq m.$$

If we take conjugates of the expression just written we have

$$\pm \bar{\pi}\bar{\tau} + \bar{\tau}\bar{\pi} = \sum A_m \cdot [(-)^{\frac{1}{2}m(m+1)} \tau\pi \pm \bar{\tau}\bar{\pi}].$$

By combining these formulas by addition and subtraction we arrive at a variety of forms.

Applying the same method to three products as factors of a single product we see that

$$\tau\pi\sigma = \sum (-)^{\frac{1}{2}m(m+1)} A_m \cdot \bar{\sigma}\bar{\pi}\tau.$$

We combine this with $\sigma\pi\tau = \sum A_m \cdot \sigma\pi\tau$ to produce various formulas. In particular let σ and τ be vectors of grade 1; then $\sigma = -\sigma$ and $\bar{\tau} = -\tau$, so that

$$\tau\pi\sigma = \sum (-)^{\frac{1}{2}m(m+1)} A_m \cdot \sigma\bar{\pi}\tau.$$

We have then

$$\sigma\pi\tau \pm \tau\pi\sigma = \sum A_m \cdot \sigma[\pi \pm (-)^{\frac{1}{2}m(m+1)} \bar{\pi}]\tau.$$

Remembering the grades of σ and τ, the only possible grades for the [] which will not give zero terms are the grades $m-2$, m, $m+2$. The bracket therefore reduces to $2\pi_m$ for the upper sign and $2(\pi_{m-2} + \pi_{m+2})$ for the lower sign. Hence

$$\sigma\pi\tau + \tau\pi\sigma = 2\sum A_m \cdot \sigma\pi_m\tau,$$
$$\sigma\pi\tau - \tau\pi\sigma = 2\sum A_m \cdot [\sigma\pi_{m-2}\tau + \sigma\pi_{m+2}\tau].$$

We finally reduce these as follows:

$$\begin{aligned}\sigma\pi\tau + \tau\pi\sigma &= 2\sum A_m \cdot \sigma[A_{m+1} \cdot \pi_m\tau + A_{m-1} \cdot \pi_m\tau] \\ &= 2\sum (-)^m A_m \cdot \sigma[A_{m+1} \cdot \tau\pi_m - A_{m-1} \cdot \tau\pi_m] \\ &= \pm 2\sum [\pi_m A_0 \cdot \sigma\tau - A_m \cdot \sigma A_{m-1} \cdot \tau\pi_m \\ &\qquad - A_m \cdot \tau A_{m-1} \cdot \sigma\pi_m],\end{aligned}$$

$$\sigma\pi\tau - \tau\pi\sigma = \pm 2\sum A_m (A_2 \cdot \sigma\tau)\ (\pi_{m-2} + \pi_{m+2}) \qquad \left(\pi \begin{matrix}\text{even}\\ \text{odd}\end{matrix}\right).$$

These formulas give the effect of transposing two vectors in such a product. They are very useful For instance when we introduce ∇ it will be seen that we may set ∇ for σ, and for $\nabla A_0 \cdot \tau(\,) + \tau A_0 \cdot \nabla(\,)$ we may write $2\phi_0(\,)$. If then we set π_m^1 for the value of π when in the vector of grade m in each term of the expansion for π_m (say $A_m \cdot \gamma_1 \cdots \gamma_m$) we write ϕ_0, thus:

$$\begin{aligned}A_m \cdot \phi_0(\gamma_1)\gamma_2 \cdots \gamma_m + A_m \cdot \gamma_1 \phi_0(\gamma_2) \cdots \gamma_m \\ + \cdots + A_m \cdot \gamma_1\gamma_2 \cdots \phi_0(\gamma_m),\end{aligned}$$

we have

$$\nabla\pi\tau + \tau\pi\nabla = \pm 2\pi A_0 \cdot \nabla\tau \mp 2\sum \pi_m^1,$$

where ∇ acts only on τ. The expression $A_0 \cdot \nabla\tau$ will be called later the *divergence* of τ. Likewise $A_2 \cdot \nabla\tau$ will be called the *curl* of τ. We see then that when curl $\tau = 0$, $\nabla\pi\tau = \tau\pi\nabla$. We cannot elaborate on these, but this single example will show how simple the transformations are without any great display of formulas of reduction.*

A few cases need to be written down, however. They will explain themselves. We use π and τ as products of grades i and j respectively, and σ as of grade 1:

$$\pi_i \tau_j = (-)^{ij}[A_{i+j} - A_{i+j-2} + A_{i+j-4} - \cdots]\tau_j \pi_i;$$

$$\tfrac{1}{2}(\pi_i \tau_j + (-)^{ij}\tau_j \pi_i) = [A_{i+j} + A_{i+j-4} + \cdots]\pi_i \tau_j;$$

$$\tfrac{1}{2}(\pi_i \tau_j - (-)^{ij}\tau_j \pi_i) = [A_{i+j-2} + A_{i+j-6} + \cdots]\pi_i \tau_j;$$

$$A_{i+j-1}\pi_i \sigma\tau_j = A_{i+j-1}[\pi_i A_{j+1} \cdot \sigma\tau_j + (A_{i-1} \cdot \pi_i \sigma)\tau_j + (-)^i \sigma A_{i+j} \cdot \pi_i \tau_j];$$

$$A_{i-j+1}\pi_i \sigma\tau_j = A_{i-j+1}[\pi_i A_{j-1} \cdot \sigma\tau_j + (A_{i-1} \cdot \pi_i \sigma)\tau_j + (-)^i \sigma A_{i+j} \cdot \pi_i \tau_j].$$

23. We consider now more particularly $A_0 \cdot \alpha_1 \cdots \alpha_{2n}$. In the first place from the development it is obvious that the form is a pfaffian. Hence if we square it we shall have a skew determinant of even order, that is, $(A_0 \cdot \alpha_1 \cdots \alpha_{2n})^2$

$$= \begin{vmatrix} 0 & A_0 \cdot \alpha_1\alpha_2 & A_0 \cdot \alpha_1\alpha_3 & \cdots & A_0 \cdot \alpha_1\alpha_{2n} \\ -A_0 \cdot \alpha_1\alpha_2 & 0 & A_0 \cdot \alpha_2\alpha_3 & \cdots & A_0 \cdot \alpha_2\alpha_{2n} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ -A_0 \cdot \alpha_1\alpha_{2n-1} & \cdots & \cdots & \cdots & A_0 \cdot \alpha_{2n-1}\alpha_{2n} \\ -A_0 \cdot \alpha_1\alpha_{2n} & \cdots & \cdots & \cdots & 0 \end{vmatrix}.$$

It is evident in this form that advancing each subscript cyclically is equivalent to moving the first row and column to the last row and column, which does not change the value of the determinant. Reversing all rows and columns does not change the determinant. The number of positive terms in the pfaffian is one greater than the number of negative terms, hence not all terms can have their signs changed without changing the value. This shows again the permutations possible in this form. If any vector is orthogonal to all the others the pfaffian vanishes. From the form $\sigma\pi\tau + \tau\pi\sigma$ if $m = 0$ we have

$$A_0 \cdot \sigma\pi\tau + A_0 \cdot \tau\pi\sigma = 2A_0 \cdot \pi_s A_0 \cdot \sigma\tau.$$

Hence if σ and τ are orthogonal the right side vanishes. Hence interchanging two orthogonal vectors in the $A_0 \cdot$ form merely changes its sign.

Let $A_0 \cdot ()$ represent $A_0 \cdot \alpha_1 \cdots \alpha_{2n}$, $A_0 \cdot (i_1, i_2, \cdots, i_{2_m})$ represent $A_0 \cdot \alpha_1 \cdots \overset{i_1, i_2, \cdots, i_{2m}}{\cdots} \cdots \alpha_{2n}$, $m < n$, and let $(-)^{i+k}A_{ik} \cdot$ be the minor of $A_0 \cdot \alpha_i \alpha_k$. It is clear that if we differentiate the expressions $A_0 \cdot ()$ as to

*By means of these forms we can reduce every expression $A_m \cdot \alpha_1 \cdots \alpha_n$ to sums and differences of products alone, thus $A_0 \cdot \alpha_1\alpha_2 = \frac{1}{2}(\alpha_1\alpha_2 + \alpha_2\alpha_1)$, $A_2 \cdot \alpha_1\alpha_2 = \frac{1}{2}(\alpha_1\alpha_2 - \alpha_2\alpha_1)$, $A_1 \cdot \alpha_1\alpha_2\alpha_3 = \frac{1}{2}(\alpha_1\alpha_2\alpha_3 + \alpha_3\alpha_2\alpha_1)$.

$A_0 \cdot \alpha_i \alpha_k$ partially, we have $(-)^{i+k-1} A_0 \cdot (i, k)$. If we differentiate the determinant in the same manner we have $2(-)^{i+k} A_{ik}$. Hence

$$A_{ik} = -A_0 \cdot () A_0 \cdot (i, k).$$

From a well known theorem in determinants if

$$\Delta(i_1, i_2, \cdots, i_m; j_1, j_2, \cdots, j_m)$$

is the determinant produced by erasing from the original determinant the rows indicated by the i subscripts, and the columns indicated by the j subscripts, then we have

$$|A_{i_s j_t}| = [A_0 \cdot ()]^{2m-2} \Delta(i_1 \cdots j_1 \cdots) \qquad (x, y = 1, 2, \cdots, m).$$

In case the subscripts i and j are the same, Δ is a skew symmetric determinant also and is the square of $A_0 \cdot (i_1, \cdots, i_m)$. The determinant on the left in any case reduces to

$$(-)^m [A_0 \cdot ()]^m \begin{vmatrix} 0 & A_0 \cdot (i_1 i_2) & \cdots & A_0 \cdot (i_1 i_m) \\ -A_0 (i_2 i_1) & 0 & \cdots & A_0 \cdot (i_2 i_m) \\ \cdot & \cdot & \cdot & \cdot \\ -A_0 (i_m i_1) & \cdots & \cdots & 0 \end{vmatrix}.$$

This vanishes if m is odd, giving a theorem as to Δ, and if m is even the determinant is the square of a pfaffian, giving a series of theorems about expressions built up from the forms $A_0 \cdot (i, j)$ in the same formal manner as the forms A_0 are built up from $A_0 \cdot \alpha_i \alpha_j$. We cite only the following:

$$A_0 \cdot (a, b) A_0 \cdot (c, d) - A_0 \cdot (a, c) A_0 \cdot (b, d) + A_0 \cdot (a, d) A_0 \cdot (b, c) = A_0 \cdot (a, b, c, d) \cdot A_0 \cdot ();$$

$$A_0 \cdot (a, b) A_0 \cdot (c, d, e, f) - \cdots = A_0 \cdot (a, b, c, d, e, f) A_0 \cdot ().$$

They are easily written and extended to all orders.

24. Space does not permit further developments here, but enough has been said to make it easy to continue. The particular forms that are related to the linear vector operator, whose general form would be

$$\phi() = \sum \alpha A_0 \cdot \beta(),$$

furnish a complete part by themselves. When the number of vector terms becomes infinite we have functional transformations, and functions of lines with related subjects enter. This development is deferred to later papers.

5. The differentiator

25. We should however mention here the differentiator (corresponding to the Hamiltonian *nabla*) related to a vector $\rho = \sum x_i \alpha_i$, expressed as depending upon n parameters, $i = 1, 2, \cdots, n$,

$$\nabla = \sum \alpha_i \partial / \partial x_i.$$

This symbol may be used anywhere just as if it were a vector, for all vector properties, and is supposed to differentiate in situ any vector dependent upon the parameters entering ρ. *It differentiates also in situ any function or operator dependent upon these parameters.* This one simple rule enables us to write out at once an indefinitely great number of forms for differential work. No special collection of formulas is necessary. The mere fact that ∇ may *follow* an operand is a principle that introduces great simplicity. We observe here that in the case of functions the process of differentiation of a function of a line used by Volterra is a case of the use of a ∇ in which the number of parameters is non-denumerably infinite.

It is evident that we have

$$d \cdot ( ) = A_0 \cdot d\rho \nabla \cdot ( )$$

for any operand dependent upon the parameters of ρ. (Note the parameters of ρ belong to the complex in which ρ is variable, and do not mean the coördinates of ρ, which would be usually more numerous and might be infinite in number.) A simple example of the use of ∇ is shown in the following.

If we have a pfaffian differential form $w = \sum y_i \, dx_i$ we know that it may be reduced to a set of terms $udv + pdq + rds + \cdots$. In case there is but one term and that is dv, the expression is called *integrable or exact.* If the single term is udv it is said to be integrable with a multiplier, u^{-1}. The number of such terms was shown by Pfaff to be not more than half the number of parameters x, if this is even, and not more than half the number plus one, if the number is odd.

Suppose we can reduce w to udv. Then we have by setting $\sigma = u \nabla v$

$$\nabla \sigma = \nabla u \nabla v + u \nabla^2 v, \qquad \text{where} \qquad \nabla^2 = \nabla \nabla,$$

whence

$$A_2 \cdot \nabla u \nabla v = A_2 \cdot \nabla \sigma, \qquad \text{and} \qquad A_3 \cdot \sigma \nabla \sigma = 0.$$

This is a necessary condition of *integrability with a multiplier.* In case $u = 1$, we have as the necessary condition $A_2 \cdot \nabla \sigma = 0$. Considering likewise the next form

$$w = udv + ds = A_0 \cdot d\rho \, (u \nabla v + \nabla s),$$

we set

$$\sigma = u \nabla v + \nabla s,$$

whence

$$\nabla \sigma = \nabla u \nabla v + u \nabla^2 v + \nabla^2 s,$$

$$A^2 \cdot \nabla \sigma = A^2 \cdot \nabla u \nabla v, \qquad A_3 \cdot \sigma \nabla \sigma = A_3 \cdot \nabla s \nabla u \nabla v, \qquad A_4 \cdot \nabla \sigma \nabla \sigma = 0.$$

This problem has been studied by Cartan* who gives the conditions above in

*Annales de l'École Normale (3), vol. 18 (1901), pp. 241–311.

scalar forms which, owing to the arbitrary differentials, are equivalent to these:

$$A_0 \cdot A_2 \nabla\sigma A_2 \cdot \delta_1\rho\delta_2\rho = 0,$$
$$A_0 \cdot A_3 \cdot \sigma\nabla\sigma A_3 \cdot \delta_1\rho\delta_2\rho\delta_3\rho = 0, \qquad \text{etc.}$$

The expression $A_0 \cdot \nabla\sigma$ is called the divergence of σ, and $A_2 \cdot \nabla\sigma$ is called the curl of σ.

We have other differential expressions, as for instance in the ordinary relativity theory in four variables, we consider $A_3 \cdot \nabla\sigma_2$ called the curl of the 2-vector (vector of grade 2) σ_2. The complement or adjunct of this in the region of order 4 is called the Lorentzian of σ. It is $A_1 \cdot A_3 \nabla\sigma_2 A_4 \omega_4$.

We may build up expressions with ∇ which are covariant for unique and invertible transformations $\phi(\rho)$. Such are

$$A_0 \cdot \nabla_2 \alpha\pi A_2 \, d_1\rho d_2\rho, \quad A_0 \cdot A_3 \cdot \pi\nabla\pi A_3 \cdot d_1\rho d_2\rho d_3\rho, \quad \cdots.$$

In fact the possibility of expressing a function in such wise implies certain covariant properties.

26. In general if we write the form $A_0 \cdot d\rho\sigma$, and consider the equation

$$A_0 \cdot d\rho\sigma = 0, \qquad \text{with} \qquad A_{2n+1} \cdot \sigma\nabla\sigma\nabla\sigma \cdots \nabla\sigma = 0 \quad \text{identically},$$

we undertake to find v such that $A_{2n} \cdot \nabla v\sigma\nabla\sigma \cdots \nabla\sigma = 0$, where both forms A are alternants. From the latter, multiplying by $d\rho$ and taking $A_{2n-1} \cdot$ we have

$$\begin{aligned} A_{2n-1}\, d\rho A_{2n} \cdot \nabla v\sigma\nabla\sigma \cdots \nabla\sigma = 0 &= dv A_{2n-1} \cdot \sigma\nabla\sigma \cdots \nabla\sigma \\ &\quad - A_0 \cdot d\rho\sigma A_{2n-1} \cdot \nabla v\nabla\sigma \cdots \nabla\sigma \\ &\quad + A_{2n-1} \nabla v\sigma d\sigma\nabla\sigma \cdots - A_0\, d\rho\sigma A_{2n-1} \nabla v\sigma\nabla\sigma \cdots + \cdots \\ &= dv A_{2n-1}\, \sigma\nabla\sigma \cdots - (n-1) A_0 \cdot d\rho\sigma A_{2n-1} \nabla v\nabla\sigma \\ &\qquad \cdots + (n-1) A_{2n-1} \nabla v d\sigma\nabla\sigma \cdots + \cdots. \end{aligned}$$

Set $dv = 0$, so that one solution is $v = \text{const.}$, and the remainder of the expression furnishes an equation which must hold for arbitrary differentials of the parameters of ρ. Setting their coefficients equal to 0, we are able to find v. For instance let us have

$$x_1 x_3\, dx_2 + x_1 x_2\, dx_3 + (x_1 + x_3 x_5)\, dx_4 + x_3 x_4\, dx_5 = 0.$$

We have $\sigma = x_1 x_3 \alpha_2 + x_1 x_2 \alpha_3 + (x_1 + x_3 x_5)\alpha_4 + x_3 x_4 \alpha_5$, the vectors $\alpha_1, \cdots, \alpha_5$ being constant. It is easy to show that $A_5 \cdot \sigma\nabla\sigma\nabla\sigma = 0$ and four variables will suffice. To find v so as to reduce the expression, we notice that since the form will be $udv + rds$

$$\sigma = u\nabla v + r\nabla s, \qquad A_2 \cdot \nabla\sigma = A_2 \cdot \nabla u\nabla v + A_2 \cdot \nabla r\nabla s,$$
$$A_3 \cdot \sigma\nabla\sigma = rA_3 \cdot \nabla s\nabla u\nabla v + uA_3 \cdot \nabla v\nabla r\nabla s.$$

Either ∇v or ∇s is common to both, hence $A_4 \cdot \nabla v\sigma\nabla\sigma = 0$. We assume the

form for v such that

$$\nabla v = v_1 \alpha_1 + v_2 \alpha_2 + v_3 \alpha_3 + v_4 \alpha_4 + v_5 \alpha_5 ,$$

and substituting all values we have a vector of grade 4 which vanishes so that each coefficient of the alternants $A_4 \cdot \alpha_1 \alpha_2 \alpha_3 \alpha_4$, $A_4 \cdot \alpha_1 \alpha_2 \alpha_3 \alpha_5$, etc., vanishes. These are here

$$\begin{aligned} x_3^2 x_5 v_3 - x_2 x_3 x_5 v_2 + x_1 x_3 x_4 v_1 &= 0 , \\ x_3^2 x_4 v_3 - x_2 x_3 x_4 v_2 + x_1 x_3 x_4 v_1 &= 0 , \\ - x_2 x_3 x_5 v_5 + x_2 x_3 x_4 v_4 - x_3 x_4 v_3 + (x_3 x_4 x_5 - x_1 - x_3 x_5) v_1 &= 0 , \\ - x_1 x_3 x_5 v_5 + x_1 x_3 x_4 v_4 + (x_3 x_4 - x_1 - x_3 x_5) v_2 &= 0 . \end{aligned}$$

From these we have

$$\begin{aligned} x_1 v_1 - x_2 v_2 + x_3 v_3 &= 0 , \\ x_4 v_2 - x_3 x_4 v_4 + x_3 x_5 v_5 &= 0 . \end{aligned}$$

These are satisfied by taking either $v = x_1/x_3$, or $x_2 x_3 + x_4$, or $x_4 x_5$.

The equation $A_0 \cdot d\rho\sigma = 0$ means that $d\rho$ is everywhere orthogonal to the congruence given by $A_2\, d\rho\sigma = 0$. This is called the *characteristic congruence*. If the equation is integrable with or without a multiplier, the congruence satisfies $A_2 \cdot \nabla\sigma = 0$ or $A_3 \cdot \sigma\nabla\sigma = 0$, and this makes it a *normal* congruence. That is, there is a function of $n - 1$ of the parameters, f, such that $A_0 \sigma \nabla f = 0$. The integral in fact is $f = \text{const}$. The other cases, $A_4 \cdot \nabla\sigma\, \nabla\sigma = 0$, etc., give special congruences worthy of study.

27. The integral $\int A_0 \cdot \sigma d\rho$ depends generally upon the *path* of ρ, that is, ρ is taken as an arbitrary function of a single parameter, s, $\rho = \rho(s)$, and $d\rho$ becomes $\rho'\, ds$. The terminal values of s furnish the limits. The integral evidently depends in general upon the form of the function $\rho(s)$. If the initial and the final value of s are the same we say we have *integrated around a loop*, and indicate the fact with the sign $\int$. If we follow any path for which we have $A_0 \cdot \sigma d\rho = 0$, it is evident that the integral from any point to any other is zero, and such paths are *orthogonal* lines for the characteristic congruence. In case they can be collected upon spreads (functions of one or more parameters) these functions are called solutions of the pfaffian form.

If we follow a loop or closed circuit in the integration, and choose a spread of order 2 (a surface), attending as usual to singularities, we may introduce elementary areas and circuits, arriving at the generalized *Stokes' Theorem*,

$$\int A_0 \cdot \sigma d\rho = \int\int A_0 \cdot A_2\, \nabla\sigma A_2 \cdot d_1 \rho d_2 \rho ,$$

over the area. This theorem still applies in case the line-integral is not taken over a circuit, if we can draw through the terminals of the actual path orthogonal lines as mentioned above so as to complete a circuit, since the integral along the lines introduced is zero. If the curl of σ is zero at all points of the area the loop integral evidently vanishes.

6. Integral invariants*

28. Suppose that we have a general function linear in $d\rho$, $Q(d\rho)$. If we integrate this over a loop the form of Stokes' theorem is

$$\int Q(d\rho) = \int\int Q'(A_1 \cdot \nabla' A_2 \, d_1\rho d_2\rho),$$

where the accent shows the operand of the accented ∇. This may be considered to be the difference between two distinct paths from an initial point to a terminal point chosen on the loop, the double integral being taken over the enclosed area. If the paths differ infinitesimally it becomes the variation of the integral for a variation in path. Therefore if

$$Q'(A_1 \nabla' A_2 [\,]) = 0 \quad \text{identically},$$

the variation is zero, and the integral is called an *integral invariant* for the variation in path. We do not necessarily have to retain the same end points, for we may move either along any line which satisfies the equation $Q(d\rho) = 0$.

When the integral is invariant it is called the integral of the *exact* differential form $Q(d\rho) = 0$. For instance, $A_2 \cdot \alpha d\rho$ is exact, the integral being $A_2 \cdot \alpha\rho$. In this case we have identically $A_2 \cdot \alpha(A_1 \nabla A_2 [\,]) = 0$, since α is constant. The lines which satisfy $A_2 \cdot \alpha d\rho = 0$ are $\rho = s\alpha$. The path could be moved parallel to α, leaving the integral invariant.

29. It happens sometimes that an integral remains invariant when the whole loop is displaced, as for instance the strength of a vortex tube. The general condition is obtained thus: let the integral be taken over a region of order n (that is, there are n independent parameters and these are varied between appropriate limits),

$$I_n = \int^n Q(A_n \cdot d_1\rho d_2\rho \cdots d_n\rho),$$

where $\int^n$ means n integral signs or an n-fold integral (the iteration of an integral in order to procure an area-integral, or a space-integral, is incidental here and may not be necessary). Suppose now that we displace the region so that ρ becomes $\rho + \delta\rho$; then we may consider the effect of this on any function of ρ by noting that

$$\delta = A_0 \cdot \delta\rho\nabla.$$

Hence

$$\begin{aligned}\delta I_n &= A_0 \cdot \delta\rho\nabla \int^n Q(A_n \cdot \delta_1\rho \cdots d_n\rho) \\ &= \int^n A_0 \cdot \delta\rho\nabla' \cdot Q'(A_n \cdots) + \int^n Q(A_n d_1\delta\rho \, d_2\rho \cdots d_n\rho) \\ &\qquad + \cdots + \int^n Q(A_n \cdot d_1\rho \cdots d_n\delta\rho) \\ &= \int^n A_0 \, \delta\rho\nabla' \cdot Q'(A_n \cdot d_1\rho \cdots d_n\rho) \\ &\qquad + \int^n Q(A_n \, d\rho' A_{n-1} \cdot \nabla' A_n \cdot d_1\rho \cdots d_n\rho).\end{aligned}$$

* The contents of this division were included in a paper presented before the Society, Chicago Section, January 2, 1913, *Integral invariants in general vector analysis.*

This is called the first form. Integrate the last term by parts, combine, and we have

$$\delta I_n = \int^n Q' (A_n \cdot \nabla' A_{n+1}\, \delta\rho\, d_1\rho \cdots d_n\rho) + \int^n [Q(A_n \cdot \delta\rho)]' A_{n-1} \cdot \nabla' A_n \cdot d_1\rho \cdots d_n\rho.$$

This is called the second form. The expansion of the last term gives a number of terms each of which is integrable once, and we have

$$\delta I_n = \int^n Q' (A_n \cdot \nabla' A_{n+1} \cdot \delta\rho\, d_1\rho \cdots d_n\rho) + \sum \int^{n-1} [(QA_n \cdot \delta\rho)_{B_i} - (QA_n \cdot \delta\rho)_{A_i}] A_{n-1} \cdot d_1\rho \cdots \overset{i}{\cdot} \cdots d_n\rho.$$

Let $\delta\rho$ follow a path so that $\delta\rho = \sigma\delta t$; then we have

$$\delta I_n/\delta t = \int^n Q' A_n \cdot \nabla' A_{n+1} \cdot \sigma d_1\rho \cdots d_n\rho + \sum \int^{n-1} [(QA_n \cdot \sigma)_{B_i} - (QA_n \cdot \sigma)_{A_i}] A_{n-1} \cdots.$$

If we had retained the first form we would have

$$\delta I_n/\delta t = \int^n A_0 \cdot \sigma\nabla' \cdot Q' (A_n \cdots) + \int^n Q (A_n \cdot \sigma A_{n-1}\, \nabla' A_n \cdots).$$

If now I_n is an absolute invariant, we must have for the paths given by σ

$$A_0 \cdot \sigma\nabla' \cdot Q' (A_n \cdots) + Q (A_n \cdot \sigma' A_{n-1} \cdot \nabla' A_n \cdots) = 0 \quad \text{identically,}$$

or for the second form

$$Q' (A_n \cdot \nabla' A_{n+1} \cdot \sigma A_n \cdots) + (QA_n \cdot \sigma)' A_{n-1} \cdot \nabla' A_n \cdots = 0 \quad \text{identically.}$$

These conditions are very general and include every form as usually stated. As an instance, consider the integral $\int A_0 \cdot \tau d\rho$. This will be an absolute invariant for the congruence σ if

$$A_0 \cdot \sigma\nabla A_0 \cdot \tau (\,) + A_0 \cdot \tau\sigma' A_0 \cdot \nabla' (\,) = 0.$$

The condition may also be written

$$\nabla' A_0 \cdot \sigma' \tau = -\tau' A_0 \nabla' \sigma \qquad \text{or} \qquad \nabla (A_0 \cdot \sigma\tau) = -A_1 \cdot \sigma A_2 \cdot \nabla\tau.$$

In case the boundary is fixed, then at the boundary $\delta\rho = 0$; if the integral is to remain invariant for all σ, we see from the second form that as the second integral is now zero, and the first holds for all σ, we must have

$$Q' (A_n \cdot \nabla' A_{n+1} \cdots) = 0 \quad \text{identically.}$$

For instance, in the example just considered we would have

$$A_0 \cdot \tau' A_1 \cdot \nabla' A_2 \cdots = 0 \quad \text{identically,}$$

whence

$$A_2 \cdot \nabla\tau = 0 \quad \text{identically.}$$

The situation considered here is the extension from an integral over two distinct paths, with fixed terminals, to an integral over two regions of order n, with fixed boundary. Such an integral is indicated by Goursat* by I_n^d. When

*Journal des Mathématiques (6), vol. 4 (1908), pp. 331.

the boundary is closed we have simply an invariant, called a relative invariant, indicated by J_n.

30. In a manner similar to that for Stokes' generalized theorem we have *Green's generalized theorem*

$$I_n = I^a_{n+1} = \mathcal{S}^n Q(A_n \cdot d_1\rho \cdots d_n\rho) = \mathcal{S}^{n+1} Q'(A_n \cdot \nabla' \cdot A_{n+1} \cdot d_1\rho \cdots d_{n+1}\rho).$$

The second integral is evidently of the type I^a. Since now a relative invariant integral may be written as an integral of order higher by unity, the condition of its being invariant may be stated. That is,

$$J_n = \mathcal{S}^n Q(A_n \cdot d_1\rho \cdots d_n\rho)$$

is a relative invariant if

$$I^a_{n+1} = \mathcal{S}^{n+1} Q'(A_n \nabla' A_{n+1} \cdot d_1\rho \cdots d_{n+1}\rho)$$

is absolute, that is to say, if

$$A_0 \cdot \sigma\nabla' \cdot Q'(A_n \cdot \nabla' A_{n+1}) + Q'(A_n \nabla' A_{n+1}\sigma'' A_n \nabla'' A_{n+1}) = 0,$$

or if

$$Q'(A_n \cdot \nabla' A_{n+1} \nabla' A_{n+2} \cdot \sigma \cdots) + (Q' A_n \cdot \nabla' A_{n+1} \cdot \sigma)'' A_n \cdot \nabla'' A_{n+1} = 0.$$

This gives us the condition (since $A_n \cdot A_2 \nabla' \nabla' A_{n+2} \cdots$ vanishes identically)

$$[Q' A_n \cdot \nabla' A_{n+1} \cdot \sigma]'' A_n \cdot \nabla'' A_{n+1} \cdots = 0.$$

For instance in the case considered above

$$A_0 \cdot \tau' A_1 \cdot \nabla' A_2 \cdot \sigma'' A_1 \cdot \nabla'' A_2 \cdot () + A_0 \cdot \tau' A_1 \cdot \nabla' A_2 \cdot \nabla' A_3 \cdot \sigma A_2 () = 0.$$

Hence, reducing, $A_2 \cdot \nabla A_1 \cdot \sigma A_2 \cdot \nabla\tau = 0$ identically.

31. Since we may always write

$$A_0 \cdot \sigma\nabla'(Q' A_n \cdot \sigma \cdots) + Q A_n \cdot \sigma A_{n-1} \sigma' A_{n-2} \nabla' A_{n-1} = A_0 \cdot \sigma\nabla' \cdot Q'(A_n \cdot \sigma \cdots) + Q(A_n \cdot \sigma' A_{n-1} \nabla' A_n \sigma),$$

the condition above becomes

$$A_0 \cdot \sigma\nabla' Q'(A_n \cdot \sigma A_{n-1}()) + Q(A_n \cdot \sigma' A_{n-1} \nabla' A_n \cdot \sigma A_{n-1}) = 0.$$

Hence if $I_n = \mathcal{S}^n Q A_n$ is an absolute invariant,

$$I_{n-1} = \mathcal{S}^{n-1} Q(A_n \cdot \sigma A_{n-1} \cdots)$$

is an absolute invariant.

32. From any integrand $Q(A_n \cdots)$ may be formed by two processes two other integrands, one by the process D, giving $Q'(A_n \cdot \nabla' A_{n+1} \cdots)$, and the other by the process E, giving $Q(A_n \cdot \sigma A_{n-1} \cdots)$. In case the latter form

vanishes identically, we mark the integral I_n of $Q(A_n \cdots)$ with an e, thus I_n^e. In any case the second application of the E process gives an identically vanishing form, hence from any I_n the E process produces an I_{n-1}^e. Integrals are of four types with respect to these processes, that is, according to whether or not they can be marked with d or e. The effect of the D process on the integrand is to mark the integral with d, of the E process to mark the integral with e. We need to prove that if an integral invariant has either character the other process will not destroy this character. That is,

$$(I^d)^e = (I^e)^d.$$

E produces from I_n^d an I_{n-1}^{de}, that is, if the first is from an exact integrand, the second will also be exact when the first is invariant, and also will be invariant. For from

$$Q'(A_n \nabla' \cdots) = 0$$

and

$$Q'(A_n \cdot \nabla' A_{n+1} \cdot \sigma \cdots) + (QA_n \cdot \sigma \cdots)' A_{n-1} \nabla' A_n \cdots = 0$$

it follows that

$$(QA_n \cdot \sigma)' A_{n-1} \cdot \nabla' \cdots = 0,$$

which is the condition that the result have an exact integrand. To be invariant we must also have

$$(QA_n \cdot \sigma A_{n-1}\sigma)' A_{n-2} \cdot \nabla' \cdots = 0,$$

which is evident. Again when $QA_n \cdot \sigma \cdots = 0$ and I_n is invariant we have also

$$Q' A_n \cdot \nabla' A_{n+1} \cdot \sigma \cdots = 0,$$

showing at once that $Q' A_n \cdot \nabla' \cdots$ is of type E. That it is also invariant demands that

$$Q' A_n \nabla' A_{n+1} \nabla' A_{n+2} \sigma \cdots + (Q' A_n \cdot \nabla' A_{n+1} \cdot \sigma)' A_{n+1} \nabla' A_n = 0,$$

which is evident. Hence the two operations are permutable when applied to an invariant, one giving $Q' \cdot A_n \nabla' A_{n+1} \cdot \sigma A_n \cdots$, the other

$$(QA_n \sigma)' A_{n-1} \nabla' A_n \cdots.$$

If Q is invariant, however, the first expression is the negative of the other. The characters are then unchanged. If we start with an invariant and apply the E process we arrive at an I^e of one lower order, usually not identically zero, and if we then apply the D process we have an integral of character I^{de} of the original order. For instance, from $A_0 \cdot \tau d\rho$ we have, by D,

$$-A_0 \cdot A_2 \nabla \tau A_2 \cdot d_1 \rho d_2 \rho,$$

by E, $A_0 \tau\sigma$. Using the other process in either case we have from the first $-A_0 \cdot A_2 \cdot \nabla\tau A_2 \cdot \sigma d\rho$, from the second $(A_0 \cdot \sigma\tau)' A_0 \cdot \nabla' d\rho$. In case now $A_0 \cdot \tau d\rho$ is an invariant the new forms are also invariant.

The D process is a sort of generalization of getting the curl, while the E process fixes certain paths for integration, along which the lower integral may be displaced.

As instances, let $I_1 = \int A_0 \cdot \tau d\rho$, and let I have both characters d, e; that means

$$A_2 \cdot \nabla\tau = 0,$$

and for a certain congruence given by $d\rho = \sigma dt$,

$$A_0 \cdot \tau\sigma = 0.$$

The expression is exact on account of character d; hence the integral is

$$u = \text{const.}$$

This is a function of ρ, and such that $\tau = z\nabla u$. Hence we have also ∇u everywhere orthogonal to the congruence line σ, by the second condition. That is, σ gives lines lying in the spread $u = c$.

Let

$$I_2^{1e} = \int\int A_2 \cdot \pi_2 A_2 \cdot d_1\rho\, d_2\rho;$$

then

$$A_0 \cdot A_2 \cdot \pi_2 A_2 \cdot \nabla A_2 = 0, \qquad A_0 \cdot A_2 \cdot \pi_2 A_2 \sigma A_1 \cdot = 0 \quad \text{identically.}$$

We have easily $A_2 \cdot \nabla\pi_2 = 0$, $A_0 \cdot A_2 \cdot \pi_2 A_2 \sigma\lambda = 0$ for every λ; that is,

$$A_1 \cdot \sigma A_2 \cdot \pi_2 = 0 \quad \text{identically.}$$

From the first condition

$$\pi_2 = A_2 \nabla\xi;$$

hence

$$A_1 \cdot \sigma A_2 \nabla\xi = 0.$$

Put $d\rho$ for σdt, and we have along the congruence

$$A_1 \cdot d\rho A_2 \nabla\xi = 0, \qquad \text{or} \qquad d\xi = \nabla' A_0 \xi'\, d\rho.$$

This means, if we set ϕ for the linear vector function $\nabla A \cdot \xi(\,)$,

$$d\xi = \phi d\rho.$$

We have then formally $d\rho = \phi^{-1} d\xi$. The complete discussion is long.

A comparison of the methods used here and those given in the coördinate treatments of these problems will show the enormous gain in simplicity.

33. To solve the differential equation

$$Q \cdot A_n \cdot d\rho\pi_{n-1} = 0$$

means that we must find a congruence of curves such that if $d\rho$ is taken along any curve the equation will be satisfied, and that $x(\rho)Q$ is an absolute invariant for the curves. That is, setting $d\rho = \sigma dt$,

$$Q \cdot A_n \cdot \sigma \cdots = 0,$$

$$(xQ)'A_n \cdot \nabla A_{n+1}\sigma \cdots + (xQA_n \cdot \sigma)'A_{n-1} \cdot \nabla' A_n \cdots = 0$$

$$= (xQ)'A_n \cdot \nabla' A_{n+1} \cdot \sigma \cdots,$$

$$0 = Q'A_n \cdot \nabla' A_{n+1} \cdot \sigma + A_0 \cdot \sigma\nabla \log x \cdot QA_n.$$

Hence we must have

$$QA_n \cdot \sigma \cdots = 0, \qquad Q'A_n \cdot \nabla' A_{n+1} \cdot \sigma \cdots = t\,QA_n \cdots,$$

for some t, and t may be zero. For instance let Q be $A_0 \cdot \xi d\rho = 0$. The condition is then

$$A_0 \cdot \sigma\xi = 0, \qquad A_0 \cdot A_2\,\nabla\xi A_2 \cdot \sigma() = tA_0 \cdot \xi().$$

For a particular case let $\xi = A_1 \cdot \rho A_2 \cdot \alpha\rho$, where α is constant. Then

$$A_2 \cdot \nabla\xi = 3A_2 \cdot \alpha\rho,\ A_0 \cdot A_2 \cdot \sigma\rho A_2 \cdot \alpha\rho = 0,\ A_1 \cdot \sigma A_2 \cdot \alpha\rho = tA_1 \cdot \rho A_2 \cdot \alpha\rho.$$

These are satisfied if we take

$$\sigma = t\rho + \text{any vector orthogonal to } \alpha \text{ and to } \rho.$$

Hence $d\rho$ must be in the plane through the origin containing the common perpendicular to α and ρ. Hence $d\rho$ is in any plane tangent to a cone of revolution about α as an axis. This cone must therefore furnish an integral, that is,

$$A_0 \cdot \rho\rho = c^2(A_0 \cdot \alpha\rho)^2.$$

34. A system of equations

$$0 = A_0 \cdot \xi_1 d\rho = A_0 \cdot \xi_2 d\rho = \cdots = A_0 \cdot \xi_n d\rho_n$$

is equivalent to a single equation

$$A_{n-1} \cdot d\rho A_n \cdot \xi_1 \cdots \xi_n = A_{n-1}\pi_n d\rho = 0.$$

The conditions then reduce to

$$A_{n-1} \cdot \pi_n \sigma = 0, \qquad A_0 \cdot \sigma\nabla A_{n-1}()\pi_n - A_0 \cdot ()\nabla A_{n-1}\sigma\pi_n = tA_{n-1}()\pi_n.$$

We cannot enlarge on the use of these forms in the study of differential equations, but enough has been said to indicate the procedures. We will, however, show the application to differential geometry.

II. DIFFERENTIAL GEOMETRY

1. Let ρ be a vector dependent upon n parameters $v_1, u_2, \cdots, u_n$, thus defining a region of order n. For instance if $n = 1$ we would call the region

a curve, if $n = 2$, a surface, etc. We have then

$$\rho = \rho(u_1, u_2, \cdots, u_n).$$

2. If now we let the parameters vary and assume that the function ρ is differentiable as to each parameter, indicating the derivative by a corresponding subscript, we have

$$d\rho = \rho_1\, du_1 + \rho_2\, du_2 + \cdots + \rho_n\, du_n$$

where $\rho_i = \partial\rho/\partial u_i$, and the n vectors are in general linearly independent.

We will indicate the adjunct of ρ_i in the region by ρ^i so that

$$\rho^i = (-)^{i-1} A_1 \cdot A_{n-1} \cdot \rho_1 \rho_2 \cdots \rho_{i-1} \rho_{i+1} \cdots \rho_n A_n \cdot \rho_1 \rho_2 \cdots \rho_n / A_0 \cdot A_n \cdot \rho_1 \cdots \rho_n A_n \cdot \rho_1 \cdots \rho_n .$$

For convenience we shall write $\nu = A_n \rho_1 \cdots \rho_n$ so that ν is an alternate vector of grade n, which is also a function of the point at which it is calculated. We also set $(I \cdot \nu)^2 = A_0 \cdot \nu\nu$, and $U\nu = \nu/I\nu$ is a unit vector. Where $\nu = 0$, we have singular points.

It is clear that

$$A_0 \cdot \rho_i \rho^j = 0, \quad \text{if} \quad i \neq j, \quad \text{and} \quad A_0 \cdot \rho_i \rho^i = 1.$$

Further

$$A_n \cdot \rho^1 \rho^2 \cdots \rho^n = \nu / A_0 \nu\nu,$$

which is seen by expanding each in terms of $\rho_1, \rho_2, \cdots, \rho_n$ and applying the theorems of determinants.

3. The expressions $A_0 \cdot \rho_i \rho_j$ for $i, j = 1, \cdots, n$ are called the *fundamental quantities* of the first order. They are commonly represented by a_{ij}. The corresponding forms $A_0 \cdot \rho^i \rho^j = a^{ij}$ are the *adjunct quantities of first order.*

It is clear that

$$|a_{ij}| = (-)^{\frac{1}{2}n(n-1)} A_0 \cdot \nu\nu, \qquad |a^{ij}| = (-1)^{\frac{1}{2}n(n-1)} A_0 \cdot \nu\nu .$$

4. The biorthogonal systems ρ_i and ρ^i are useful. Suppose for instance we have a system of values l_s which we attach to ρ_s giving a vector

$$\lambda = \sum \rho_s l_s;$$

then we may define a new system of values

$$m = A_0 \cdot \rho^r \lambda,$$

and from the properties of the vectors ρ we have

$$\mu = \sum m_r \rho_r = \sum \rho_r A_0 \rho^r \lambda = \lambda .$$

The coefficients m and the coefficients l therefore give the same vector when attached to the proper vectors. For the vector ρ in terms of ρ_i, and the vector

ρ in terms of ρ^i we may introduce a new system of n independent vectors defining the region, and their biorthogonal system. As we are dealing with the vectors themselves we shall not be concerned with their mode of expression. Such questions however play an important part in the developments of Ricci (*Lezioni sulla Teoria delle Superficie*) and those who follow him.

5. If we have a vector $\xi = \sum X_r \rho^r$, and in the other form, $\xi = \sum X^r \rho_r$, then $\xi/I\rho_r$ is such that $X_r/I\rho_r =$ the orthogonal projection of ξ on ρ_r. In other words the orthogonal projections are given by $A_0 U\rho_r \xi$. Similarly for the other form.

6. If we differentiate $A_0 \rho_r \rho^s = 0$ or 1 we have in either case

$$A_0 \rho_{rt} \rho^s + A_0 \rho_r (\rho^s)_t = 0,$$

a useful formula. From it we have the Christoffel * form

$$\left\{ \begin{matrix} rs \\ t \end{matrix} \right\} = A_0 \rho_{rs} \rho^t = - A_0 \rho_r (\rho^t)_s = - A_0 \rho_s (\rho^t)_r .$$

Evidently

$$\begin{aligned} \rho_{rs} &= - \sum \rho_t A_0 \cdot \rho_r (\rho^t)_s + A_1 \cdot U\nu A_{n+1} \cdot U\nu\rho_{rs} \\ &= - \sum \rho_t A_0 \cdot \rho_s (\rho^t)_r + A_1 \cdot U\nu A_{n+1} \cdot U\nu\rho_{rs} . \end{aligned}$$

The second term† in each form gives that part of ρ_{rs} outside the region of order n. The other Christoffel form is

$$\left[\begin{matrix} rs \\ t \end{matrix} \right] = A_0 \cdot \rho_{rs} \rho_t .$$

7. Since $A_{n+1} \cdot \rho_1 \rho_2 \cdots \rho_n \rho^m = 0$ we have

$$dA_{n+1} \cdot \nu\rho^m = 0, \qquad \text{or} \qquad A_{n+1} \cdot \nu (\rho^m)_s = - A_{n+1} \cdot \nu_s \rho^m ,$$

and likewise

$$A_{n+1} \cdot \nu\rho_{rs} = - A_{n+1} \cdot \nu_r \rho_s = - A_{n+1} \cdot \nu_s \rho_r .$$

There is usually a non-vanishing expression of first degree which is orthogonal to the region under consideration $(\rho_1, \cdots, \rho_n)$

$$A_1 \cdot \nu A_{n+1} \cdot \nu\rho_{rs} = - A_1 \cdot \nu A_{n+1} \cdot \nu_s \rho_r = - A_1 \cdot \nu A_{n+1} \cdot \nu_r \rho_s .$$

We note that $U\nu$ might as well have been used here as ν. We can construct n^4 expressions

$$A_0 \cdot A_{n+1} \cdot \rho_{rs} \nu A_{n+1} \nu\rho_{tu} ,$$

which can also be written

$$A_0 \cdot \rho_r \nu_s A_{n+1} \cdot \nu_t \rho_u , \text{ or } A_0 \cdot \rho_r A_1 \cdot \nu_s A_{n+1} \cdot \nu_t \rho_u , \text{ or } A_0 \cdot \rho_u \nu_t A_{n+1} \cdot \rho_r \nu_s ,$$

* This is easily reduced to the usual cartesian form by writing out ρ_{rs} in terms of $\rho_1, \cdots, \rho_n$. For the other Christoffel form expand ρ_{rs} in terms of $\rho', \cdots, \rho^n$.

† It is to be noted that in general ρ_{rs} has components not linearly expressible in terms of $\rho_1, \cdots, \rho_n$, as for instance for a surface, which is imbedded in a three-dimensional space.

and in other forms produced by permissible interchanges of the subscripts. However this is a quadratic in ρ_r, ρ_u of the form $A_0 \rho_r \phi \rho_u$, where ϕ is a linear vector function. Its transverse would be

$$A_0 \cdot \rho_r \nu_t A_{n+1} \nu_s \rho_u .$$

We will set

$$A_1 \cdot (U\nu)_s A_{n+1} \cdot (U\nu)_t ( ) = \mathrm{N}_{st} ( ) ,$$
$$A_1 \cdot (U\nu)_t A_{n+1} \cdot (U\nu)_s ( ) = \bar{\mathrm{N}}_{st} = \mathrm{N}_{ts} .$$

The difference of these two forms has received a symbol due to Riemann:

$$(ru, st) = A_0 \cdot \rho_r [(U\nu)_s A_{n+1} \cdot (U\nu)_t - (U\nu)_t A_{n+1} \cdot (U\nu)_s] \rho_u = A_0 \cdot \rho_r (\epsilon \mathrm{N}_{st}) \rho_u .$$

There is also another Riemann symbol * of order four,

$$\{ru, st\} = A_0 \cdot \rho_r \mathrm{N}_{st} \rho^u - A_0 \cdot \rho_r \mathrm{N}_{ts} \rho^u = A_0 \rho_r (\epsilon \mathrm{N}_{st}) \rho^u ,$$

where $\epsilon(\mathrm{N}) = \frac{1}{2}(\mathrm{N} - \bar{\mathrm{N}})$, $\bar{\mathrm{N}}$ being the transverse linear vector function of N. A better form may be given these by utilizing a previous formula,† which gives

$$\{ru, st\} = A_0 A_2 \rho_r \rho_u A_2 \cdot (U\nu)_s (U\nu)_t = - A_0 \cdot \rho_{rs} U\nu A_0 \cdot \rho_{ut} U\nu + A_0 \cdot \rho_{rt} U\nu A_0 \cdot \rho_{us} U\nu ;$$
$$\{ru, st\} = A_0 A_2 \rho_r \rho^u A_2 \cdot (U\nu)_s (U\nu)_t = - A_0 \cdot \rho_{rs} U\nu A_0 \cdot (\rho^u)_t U\nu + A_0 \cdot \rho_{rt} U\nu A_0 \cdot (\rho^u)_s U\nu .$$

From these expressions it is evident that the two symbols in each couple may be interchanged by changing the sign, and the two couples may be interchanged.

8. The expression $A_0 \cdot d\rho d\rho = \sum A_0 \cdot \rho_r \rho_s du_r du_s$ is called *the first fundamental form*. In case the Riemann symbols vanish it can be reduced to the sum of not more than n squares, that is, new variables may be found in terms of which the parameters u may be expressed so that, for the new derived vectors ρ', we will have

$$A_0 \cdot d\rho d\rho = \sum A_0 \rho'_r \rho'_r (du'_r)^2 .$$

The condition is both necessary and sufficient. If some or all of the Riemann symbols do not vanish, then it will be necessary to express $A_0 \cdot d\rho d\rho$ as the sum of not more than $\frac{1}{2}n(n+1)$ squares. From one point of view this is equivalent to considering the region to be embedded in a region of not more than $\frac{1}{2}n(n+1)$ dimensions in which we may take the directions of ρ_i as orthogonal.

For instance a curve on an ordinary surface has a normal in the surface and

* These may be identified with the forms commonly used by substituting and reducing.

† $A_0 A_2 \rho_r \rho_u A_2 (U\nu)_s (U\nu)_t = + A_0 \rho_r (U\nu)_t A_0 \rho_u (U\nu)_s - A_0 \rho_r (U\nu)_s A_0 \rho_u (U\nu)_t .$

also has the normal to the surface as a normal, or indeed all lines in the plane of these two. Hence the ds^2 is expressible in terms of three differentials in ordinary space which encloses the surface, in the form

$$ds^2 = dx^2 + dy^2 + dz^2.$$

On the surface alone it would usually have terms of the form

$$du^2, \quad dudv, \quad dv^2,$$

and would not be reducible to the sum of two squares.

9. We may improve the forms we have reached so far by the introduction of the operator which generalizes the Hamilton ∇. On account of this generalization we shall use ∇ in a general sense. It is defined thus:

$$\nabla = \sum \rho^i \, \partial / \partial u_i .$$

It follows that $d(\,) = A_0 \, d\rho \nabla \cdot (\,)$ for any function of the parameters u only. Hence we will have $\partial / \partial u_r = A_0 \cdot \rho_r \, \nabla \cdot$; therefore we may write for $(Uv)_r$

$$(Uv)_r = A_0 \rho_r \, \nabla \cdot Uv .$$

The operator

$$A_0 (\,) \nabla \cdot Uv = \Phi$$

plays a very important part in all these problems and gives of course the vector rate of change of Uv in the direction given by the unit vector operated upon. With this symbol we may write the Riemann symbols in the forms

$$(ru, st) = A_0 A_2 \rho_r \rho_u A_2 \Phi \rho_s \Phi \rho_t$$

with the understanding that s, t may be interchanged with r, u;

$$\{ru, st\} = A_0 A_2 \rho_r \rho^u A_2 \Phi \rho_s \Phi \rho_t .$$

The number of Riemann symbols with two pairs of identical subscripts is $\frac{1}{2} n (n-1)$, with one pair of similar indices $\frac{1}{2}(n-1)(n-2)$, and with all different $\frac{1}{12} n (n-1)(n-2)(n-3)$, giving a total of $\frac{1}{12} n^2 (n^2 - 1)$. It is clear that for any four directions in the region we may have a symbol $A_0 A_2 \alpha\beta A_2 \Phi\gamma\Phi\delta$, which for unit vectors may be called the Riemann symbol for those directions. There is also, if we take differentials for the intensities of the vectors, a differential quadrilinear expression

$$G_4 = A_0 \cdot A_2 \cdot d' \rho \, d'' \rho \, A_2 \cdot \Phi \, d''' \rho \, \Phi \, d'''' \rho .$$

Since every symbol is invariant for a change of parameters, this is a covariant.

We see at once that for any six directions we have a similar form

$$G_6 = A_0 \cdot d_1 \rho \, d_2 \rho \, d_3 \rho \, A_3 \cdot \Phi \, d_4 \rho \, \Phi \, d_5 \rho \, \Phi \, d_6 \rho ,$$

and there will be higher forms as high as G_{2n}. These are all covariants, and are generalizations of the Christoffel quadrilinear covariant G_4.

10. From these we derive covariant quotients, which give us in the first instance the Riemann curvature for the plane of $d_1\rho$, $d_2\rho$, and by generalizations, Riemann curvatures for the regions of $d_1\rho$, $d_2\rho$, $\cdots$. These are

$$K_m(d_1 \cdots d_m) = \frac{A_0 \cdot A_m \cdot d_1\rho d_2\rho \cdots d_m\rho A_m \Phi d_1\rho \Phi d_2\rho \cdots \Phi d_m\rho}{A_0 \cdot A_m \cdot d_1\rho d_2\rho \cdots d_m\rho A_m \cdot d_1\rho d_2\rho \cdots d_m\rho}.$$

When we arrive at the covariant of order K_n it is evidently the single coefficient of the Christoffel form for the same order. This number is called the Kronecker-Gaussian curvature of the space.

11. If we construct the second differential of ρ, $d^2\rho$, we have two parts to it, namely, that part which has components in the region, and the part which is orthogonal to the region. Since we have identically

$$A_{n+1} \cdot U\nu d\rho = 0,$$

it follows that

$$A_{n+1} \cdot dU\nu d\rho = -A_{n+1} \cdot U\nu d^2\rho.$$

But

$$d^2\rho = \sum \rho_i A_0 \cdot \rho^i d^2\rho + A_1 \cdot U\nu A_{n+1} \cdot U\nu d^2\rho$$
$$= \sum \rho_i A_0 \cdot \rho^i d^2\rho - A_1 \cdot U\nu A_{n+1} \cdot dU\nu d\rho.$$

We set now

$$\xi = A_1 \cdot U\nu A_{n+1} \cdot dU\nu d\rho.$$

It is evident that $A_0 \cdot \rho_i \xi = 0$ for $i = 1, \cdots, n$, so that ξ is not in the region. ξ is called the *vector second differential form* since the coefficients of its expansion are scalar second differential forms.* We evidently have

$$\xi = \sum A_1 \cdot U\nu A_{n+1} \cdot (U\nu)_r \rho_s \, du_r \, du_s.$$

The coefficients $A_1 \cdot U\nu A_{n+1} (U\nu)_r \rho_s$ are fundamental vectors of second order. They depend upon the parameters and are not covariants, but they furnish a number of covariants. We see that

$$A_{n+1} \cdot dU\nu d\rho = -A_{n+1} \cdot U\nu\xi.$$

Also

$$A_1 \cdot U\nu A_{n+1} \cdot dU\nu d\rho = -A_1 \cdot U\nu A_{n+1} \cdot U\nu\xi.$$

But in general

$$A_1 \cdot A_n \, \alpha_1 \alpha_2 \cdots \alpha_n \, A_{n+1} \gamma\beta_1\beta_2 \cdots \beta_n$$
$$= (-)^n \gamma A_0 \cdot A_n \cdot \alpha A_n \beta + (-)^n A_1 \cdot (A_{n-1} \cdot \gamma A_n \cdot \alpha) \cdot A_n \beta,$$

hence

$$A_1 \cdot U\nu A_{n+1} \cdot dU\nu d\rho = d\rho A_0 \cdot U\nu dU\nu + A_1 \cdot (A_{n-1} \cdot d\rho A_n \cdot U\nu) A_n \cdot dU\nu$$
$$= A_1 \cdot (A_{n-1} \cdot d\rho A_n \cdot U\nu) A_n \cdot dU\nu,$$

* Cf. Wilson and Moore, Proceedings of the American Academy of Arts and Sciences, vol. 52 (1916), pp. 270–368.

and also

$$A_1 \cdot U\nu A_{n+1} \cdot U\nu\xi = \xi + A_1 \cdot (A_{n-1} \cdot \xi U\nu) A_n \cdot U\nu .$$

But ξ is orthogonal to all vectors ρ, hence $A_{n-1} \cdot \xi U\nu = 0$. Hence

$$\xi = A_1 \cdot (A_{n-1}\, d\rho U\nu) A_n \cdot dU\nu .$$

If we go back to the theory of curves we see that if we write the differential arc in the usual form $d\rho = ds \cdot \alpha$,* then $d^2\rho = d^2 s\alpha + ds d\alpha$ and as $A_0\, \alpha d\alpha = 0$ it follows that the part of $d^2\rho$ outside the region consists of part of the term containing $d\alpha$. If $d\alpha = c\beta ds$, $\xi = A_1 \cdot A_n \cdot U\nu A_{n+1} \cdot U\nu\beta \cdot c \cdot ds^2$, which is the part of $\beta \cdot c \cdot ds^2$ orthogonal to every vector $\rho_1, \cdots, \rho_n$.

12. It is evident that the first differential quadratic and whatever is derivable from it of a covariantive or invariantive character will be true of a region which is flexible but inextensible. That is to say, bending the region would have no effect upon the first fundamental form nor upon the quantities derived from it without differentiation. However the second fundamental vector form does not possess this character and demands a rigid space. This is due to the fact that $d^2\rho$ contains a part which is not in the region itself. For instance if we are dealing with a curve in three or more dimensions, the first fundamental form is merely ds^2, and as long as we do not consider $d^2\rho$ we merely work in the line itself which for all such purposes might as well be straight. If however we consider $d^2\rho = d^2 s \cdot \alpha + ds \cdot d\alpha$ where α is the unit tangent (or better the hypernumber belonging to that particular tangent), then we must either abandon the line or consider $d\alpha = 0$ which makes the line really straight.

It is to be noticed that the expressions (Christoffel symbols) $A_0 \cdot \rho_{rs}\, \rho_t$ can be found from the region itself, since

$$2A_0\, \rho_{rs}\, \rho_t = (A_0\, \rho_r\, \rho_t)_s + (A_0\, \rho_s\, \rho_t)_r - (A_0\, \rho_r\, \rho_s)_t ,$$

all of which could be computed in a practical case from measurements inside the region. From these we can compute the expressions $A_0 \cdot \rho_{rs}\, \rho^t$. The Riemann symbols are also computable in the same manner, since

$$\begin{aligned}(rt,\, su) &= (A_0 \cdot \rho_{rs}\, \rho_t)_u - (A_0 \cdot \rho_{ru}\, \rho_t)_s \\ &\quad + \sum [A_0 \cdot \rho^h \rho^k (A_0 \cdot \rho_{ru}\, \rho_h\, A_0 \cdot \rho_{st}\, \rho_k - A_0 \cdot \rho_{rs}\, \rho_h\, A_0 \cdot \rho_{tu}\, \rho_k)] .\end{aligned}$$

Covariants built out of these forms can likewise be computed directly in the space, and would be independent of any particular directions such as r, s, t, u.

13. If $\alpha = Ud\rho$, then $d\alpha$ has two components, one in the region, and one exterior to the region. The latter is ξ/ds. The former is called the geodesic vector curvature of the curve of which α is the tangent. If it vanishes the curve is called a geodesic for the region. This may also be found as a curve of

* Where α is the unit tangent vector.

minimal length. For let

$$s = \int \sqrt{A_0 \cdot d\rho d\rho};$$

then

$$\delta s = \int \delta \sqrt{A_0 \, d\rho d\rho} = \int A_0 \, d\delta\rho d\rho / ds = \int A_0 \, d\delta\rho\alpha .$$

Integrating by parts, the first part being then exact and vanishing for an extremal, we have as the condition of a minimum (or maximum) $A_0 \, \delta\rho d\alpha = 0$ for every $\delta\rho$ in the space. That is to say, we have $d\alpha$ wholly exterior to the space. We might also write this fact in the form

$$A_{n-1} \, d\alpha A_n \, U\nu = 0 .$$

It is understood here of course that the $d\alpha$ is for a displacement along the curve.

14. The covariant G_4 can be written

$$G_4 = A_0 \cdot A_2 \cdot d_1\rho d_2\rho A_2 \cdot d_3 U\nu d_4 U\nu$$

and is thus dependent upon the four directions of differentiation. Now we have

$$dU\nu = d(\nu/I\nu) = d\nu/I\nu - U\nu A_0 \, U\nu d\nu/I\nu .$$

Hence as $A_2 \, A_n \, \alpha_1 \cdots \alpha_n \, A_n \, \alpha_1 \cdots \alpha_n = 0$,* and therefore $A_2 \cdot U\nu U\nu = 0$,

$$G_4 = A_0 \cdot A_2 \cdot d_1\rho \, d_2\rho \, [A_2 \cdot d_3\nu \, d_4\nu - A_2 \cdot U\nu \, (d_3\nu A_0 \cdot d_4\nu - d_4\nu \cdot A_0 \cdot d_3\nu) \, U\nu]/A_0 \cdot \nu\nu .$$

If we let four tangents to the space, that is, four unit directions, be α_1, α_2, α_3, α_4, we have

$$K_4 = A_0 \cdot A_2 \cdot \alpha_1 \alpha_2 \, A_2 \cdot \Phi \, \alpha_3 \Phi \, \alpha_4 / A_0 \cdot A_2 \cdot \alpha_1 \alpha_2 \, A_2 \cdot \alpha_3 \alpha_4 .$$

The Riemann symbol (rs, tu) may be computed from observations in the space as follows:

We find first

$$(A_0 \cdot \rho_{rt} \rho_s)_u - (A_0 \cdot \rho_{ru} \rho_s)_t .$$

Then we find in the region $A_0 \rho'_{rt} \rho'_{su} - A_0 \cdot \rho'_{ru} \rho'_{st}$ which is subtracted from the other value (§ 12). The difference is (rs, tu). The accents on ρ indicate that we are measuring so much of ρ_{rt}, etc., as we observe in the region in question. In the first expression these accents are not needed, as the form shows that the other parts of the second derivatives drop out anyhow. If we multiply by du_u and du_t and sum for all values of the differentials we have

$$d_1 A_0 \cdot d_2 \rho_r \, \rho_s - d_2 A_0 \cdot d_1 \rho_r \, \rho_s - A_0 \cdot d_2 \rho'_r \, d_1 \rho'_s + A_0 \cdot d_1 \rho'_r \, d_2 \rho'_s .$$

* $A_2 A_n \alpha_1 \cdots \alpha_n A_n \alpha_1 \cdots \alpha_n = \sum_{i,j} A_0 A_{n-1} \alpha_1 \cdots \alpha_n A_{n-1} \alpha_1 \cdots \alpha_n A_2 \alpha_i \alpha_j$, which will contain $(A_2 \alpha_i \alpha_j + A_2 \alpha_j \alpha_i) A_0 A_{n-1} \alpha_1 \cdots \alpha_n A_{n-1} \alpha_1 \cdots \alpha_n$, and thus all terms cancel.

Multiply again by $du_s\, du_r$ and sum again, and we have

$$G_4 = d_1 A_0 \cdot d_2 d_3 \rho d_4 \rho - d_2 A_0 \cdot d_1 d_3 \rho d_4 \rho - A_0 \cdot d_3 d_2 \rho' d_1 d_4 \rho' + A_0 \cdot d_1 d_3 \rho' d_2 d_4 \rho' .$$

This is the best form for this quadrilinear covariant, since by choosing the directions of the four differentials properly we may obtain any one of the Riemann forms.

15. The Riemann-Christoffel tensor G_4'' can be written in the form

$$(A_0 \cdot A_2 \cdot \alpha\beta A_2 \cdot \nabla_1 \nabla_2 A_0 \cdot \gamma\delta)'' ,$$

where ∇_1 acts on γ only, ∇_2 on δ only, α, β are purely arbitrary, the double accent means that only components exterior to the space are concerned. $\alpha, \beta, \gamma, \delta$ are unit vectors. If this is to vanish for any α, β, then

$$(A_2 \cdot \nabla_1 \nabla_2 A_0 \cdot \gamma\delta)'' = 0 ,$$

and this may be written

$$(A_2 \cdot \nabla\phi_\gamma' \delta)'' = 0 ,$$

where ϕ_γ is the derivative dyad of γ, or $\phi_\gamma \zeta dt = d_\zeta \gamma$, the differential of γ in the ζ direction. In other terms $\phi_\gamma = A_0 \cdot ( ) \nabla \cdot \gamma$ and $\phi_\gamma' = \nabla A_0 \cdot \gamma ( )$.

If this vanishes for every δ we must have $(A_2 \cdot \nabla\gamma)'' = 0$. This can vanish for every γ only when all the Riemann symbols vanish, which interprets the meaning of the vanishing of all these symbols.

16. The Einstein gravitation equations translated into our notation are

$$\begin{aligned} G_{ij} &= -(\sum A_0 \cdot \rho'_{ij} \rho^k)_k - \sum A_0 \cdot \rho'_{ik} (\rho^k)_j \\ &\qquad + (A_0 \cdot \nabla\rho_i)_j + \sum A_0 \cdot \rho'_{ij} (\rho^k)_k \\ &= -A_0 \nabla\rho'_{ij} - \sum A_0 \cdot \rho'_{ij} (\rho^k)_k - \sum A_0 \cdot \rho'_{ik} (\rho^k)_j \\ &\qquad + (A_0 \cdot \nabla\rho_i)_j + \sum A_0 \cdot \rho'_{ij} (\rho^k)_k \\ &= -A_0 \cdot \nabla\rho'_{ij} - \sum A_0 \cdot \rho'_{ik} (\rho^k)_j + (\sum A_0 \rho^k \rho_{ik})_j \\ &= -A_0 \cdot \nabla\rho'_{ij} - \sum A_0 \cdot \rho'_{ik} (\rho^k)_j + \sum A_0 \cdot \rho'_{ik} (\rho^k)_j + \sum A_0 \cdot \rho^k \rho_{ijk} \\ &= -A_0 \cdot \nabla\rho'_{ij} + A_0 \cdot \nabla\rho_{ij} = (A_0 \cdot \nabla\rho_{ij})'' = 0 \text{ for all } i, j . \end{aligned}$$

Therefore the significance is easily seen to be that the convergence of the exterior parts of the second differentials is zero. That is, if α is a unit tangent, $A_0 \cdot \nabla d\alpha'' = 0$. Hence the sum of the normal curvatures of n orthogonal geodesics at the point is zero. The Einstein curvature is

$$\sum g^{ij} G_{ij} = \sum A_0 \cdot \rho^t \nabla^2 \rho_t'' = \sum_{rs} A_0 \cdot \rho^r \rho^s A_0 \cdot \rho^t \rho_{trs} .$$

The Riemann-Christoffel covariant is written also without the accents (since components in the space lead to terms that cancel)

$$A_0 \cdot A_2 \cdot \alpha\beta A_2 \cdot \nabla_1 \nabla_2 A_0 \cdot \gamma\delta ,$$

and if we substitute for γ the adjunct of β, and sum for n linearly independent values β_i with β^i, we have the Einstein covariant in the form

$$G_{\alpha\delta} = \sum A_0 A_2 \alpha\beta_i A_2 \nabla_1 \nabla_2 A_0 \beta^i \delta = 0 \quad \text{for all } \alpha, \delta .$$

This gives $\sum A_0 A_2 \alpha\beta_i A_2 \nabla\phi'_\delta \beta^i$ where ∇ operates only on the β^i. This takes the form

$$\sum (A_0 \cdot \phi_\delta \alpha\beta^i A_0 \beta_i \nabla - A_0 \alpha\nabla A_0 \beta^i \phi_\delta \beta_i) ,$$

which is another form of the potential.

If we start from the forms $(A_0 \cdot \rho_{tr} \rho^s)''_u - (A_0 \cdot \rho_{ru} \rho^s)''_t$ and, making $s = u$, sum, we have

$$G_{rt} = A_0 \cdot \nabla\rho_{rt} + \sum A_0 \rho_{rt} (\rho^s)_s = (A_0 \nabla\rho_r)_t = A_0 \cdot \nabla\rho''_{rt}$$

as before.

The simplest way to get the second derivatives is to choose paths in the region on the geodesics. There is then no normal in the region. If then the convergence in the region of these geodesic normals vanishes, the space is of the gravitational character.

If the Riemann-Christoffel covariants vanish for an ordinary space we know it is a plane, and a surface nabla of the normals (which would run out into three-dimensional space) all vanish. To be an Einstein space of two dimensions it would be necessary for the convergence in the space of the geodetic normals to vanish. For a sphere, for instance, the normals all pass through the center, their differentials have any direction on the surface, and would not necessarily give a convergence equal to zero. A sphere is therefore not a possible Einstein surface.

Tensors

17. As commonly used the term tensor means merely a multiplex of coefficients obtained when we substitute in a lineo-linear function of m variables the various vectors chosen m at a time from the set $\rho_1, \cdots, \rho_n$. We shall use Ξ to indicate a tensor, which from our point of view is an operator. To indicate the operands we will insert m parentheses after Ξ. A tensor is an extension of the term dyad, triad, etc., used by Gibbs. We will write then a tensor of order 1 as $\Xi()$, one of order 2 as $\Xi()()$, etc. The first fundamental form is a tensor since it has the form $A_0 \cdot ()()$. From it we derive the set of quantities $A_0 \cdot \rho_r \rho_s$. It is clear that all our A symbols, $A_m, \cdots$ furnish tensors. A tensor is a *covariant* tensor when it reduces to a form independent of $\rho_1, \cdots, \rho_n$. All the symbols A and their combinations furnish tensors that are covariant.

A tensor is a *derivative* tensor, and of order higher by unity than another, when it is formed from the other by the process

$$\Xi()() \cdots () A_0 \cdot \nabla() .$$

For instance from the tensor ξ (a covariant, as $\rho_1, \cdots, \rho_n$ are not in evidence) we derive $\xi A_0 \nabla()$ where ∇ operates only on ξ. If $\xi = \nabla X$ then we have $\nabla X A_0 \nabla()$. If ξ is a differential operator, as $A_0 \cdot () \nabla \cdot \sigma$, then we derive

$$A_0 \cdot () \nabla \cdot \sigma \cdot A_0 \cdot \nabla().$$

If we indicate direction by a subscript as σ_r, then the derived tensor demands that $\sigma_{rs} = \sigma_{sr}$. This condition is necessary. It is further a sufficient condition that Ξ be a derived or differential tensor.

All differentials or derivatives taken in this manner without reference to the parameters u are called covariant differentials. The so-called contravariant systems would be found by using the system ρ^i.

18. To make translations from the ordinary notations we have to notice that the covariant tensor A_{xy} means the form $\sum A_{xy} A_0 \cdot \rho^x() A_0 \cdot \rho^y()$, the contravariant tensor A^{xy} means $\sum A^{xy} A_0 \cdot \rho_x() A_0 \cdot \rho_y()$, and the mixed tensor A^y_x means $\sum A^y_x A_0 \cdot \rho^x() A_0 \cdot \rho_y()$.

The so-called product of two tensors of orders a, b is a tensor of order $a + b$. For instance the product $A_{uv} B^r_s$ means

$$\sum A_{uv} B^r_s A_0 \cdot \rho^u() A_0 \cdot \rho^v() A_0 \cdot \rho^r() A_0 \cdot \rho_s().$$

In expressions where "dummy" indices appear, that is to say, duplicated indices, the expressions are summed as to the operational terms indicated, for instance, A^s_{urs} means $\sum_s A^s_{urs} A_0 \cdot \rho^u() A_0 \cdot \rho^r() A_0 \cdot \rho^s() A_0 \cdot \rho_s()$.

We will translate some of the current notations into these forms:

$$a_{uv,w} = [uv, w] = A_0 \cdot \rho_{uv} \rho_w, \qquad \{uv, w\} = A_0 \cdot \rho_{uv} \rho^w,$$
$$A^i = \sum A^i A_0 \cdot \rho^i(), \qquad (A^i)_j = \sum A^i_1 A_0 \cdot \rho^i() A_0 \cdot \nabla_1(),$$

where the subscript 1 shows the operand of ∇. This is covariantive differentiation.

With these indications it is easy to reduce any expression into the vector forms, and then to reduce the vector forms to compact expressions which from their forms alone show that they are covariantive.

Curves

19. Let α be the unit tangent to a curve (that is, a function of one parameter s), which gives $A_0 \cdot \alpha\alpha = 1$. We have therefore at once

$$A_0 \cdot \alpha \, d\alpha = 0.$$

Hence $d\alpha$ is orthogonal to α. Let β be a unit vector so that

$$d\alpha = c'_1 \beta \, ds.$$

c'_1 is called the curvature and β the principal normal. Usually part of β is

in the space of the curve, part not in the space. For instance, a curve on a sphere may have its principal normal lying in the embedding three-dimensional space. We may also write, then,

$$d\alpha = (g_1 \lambda + n_1 \nu)\, ds,$$

where now g_1 is the geodesic curvature, n_1 is the normal curvature, λ is a unit normal in the space, and ν is a unit normal out of the space.

Since $A_0 \cdot \alpha\beta = 0$, differentiating again, $A_0 \cdot \alpha d\beta = -A_0 \cdot \beta d\alpha = -c_1'$. Hence we have for $d\beta$ (the $d\rho$ is still along the curve)

$$d\beta = -c_1' \alpha + c_1'' \gamma,$$

where γ is called the binormal, and c_1'' is the second curvature. Since the first term gives the entire projection on α we must have $A_0 \cdot \alpha\gamma = 0$, and since $A_0 \cdot \beta d\beta = 0$, we must have also $A_0 \cdot \beta\gamma = 0$. Continuing we find the trinormal, etc., and the third, etc., curvatures, giving the generalized Frenet equations:

$$\begin{aligned} d\alpha &= c_1' \beta ds, \\ d\beta &= (c_1'' \gamma - c_1' \alpha)\, ds, \\ d\gamma &= (c_1''' \delta - c_1'' \beta)\, ds, \\ &\cdots\cdots\cdots \\ d\mu &= -c_1^{(m)}{}_{K}\, ds. \end{aligned}$$

When $c_1' = 0$ the other curvatures become indefinite but are taken as zero by convention. If $c_1' \neq 0$ and $c_1'' = 0$, from the equations above

$$A_3 \cdot d\rho d^2\rho d^3\rho = 0.$$

The condition is necessary and sufficient. Hence, differentiating,

$$A_3 \cdot d\rho d^2\rho d^4\rho = 0 = A_3 \cdot d\rho d^2\rho d^5\rho,$$

and all higher curvatures may be taken as zero.

20. A congruence of curves is defined by the equation

$$A_2 \cdot \lambda d\rho = 0,$$

where λ is a given unit vector in the region.* Let us consider n independent congruences, one curve of each congruence through any one point, and at the point let all the congruences be mutually orthogonal. To indicate the orthogonality we shall use α instead of λ. Then

$$A_0 \cdot \alpha_i \alpha_h = \delta_{ih},$$

and using ϕ_{α_i} to indicate $A_0 \cdot (\,) \nabla \cdot \alpha_i$, the operator that gives the differential

* The notation of Ricci is as follows in vector forms: $\lambda_{ij\rho\tau} = A_0 \cdot \lambda_i U_{\rho\tau}$, $\lambda_{ih}^{(\rho)} = A_0 \cdot \lambda_h \rho^\tau$, where i and h define different congruences; $\lambda_{ij\rho\tau s} = A_0 \cdot \rho_\tau (\lambda_i)_s$, $\gamma_{ijk} = A_0 \cdot \lambda_j (\lambda_i)_k$.

of α_i in the direction of the unit operand, we have

$$A_0 \cdot \alpha_i \phi_{\alpha_j} \alpha_k = - A_0 \cdot \alpha_j \phi_{\alpha_i} \alpha_k \quad \text{for all} \quad i \neq j \neq k .$$

Such an expression is called a *coefficient of rotation* for a reason that will appear. There are $\frac{1}{2}n^2(n-1)$ such expressions. We have also

$$A_0 \cdot \alpha_h \phi_{\alpha_h} \alpha_l = 0 \quad \text{if} \quad h \neq l .$$

We may expand $A_0 \cdot \rho_r \phi_{\alpha_h} \rho_s$ in terms of $\sum A_0 \cdot \alpha_i \rho_r A_0 \cdot \alpha_j \rho_s A_0 \cdot \alpha_i \phi_{\alpha_h} \alpha_j$. This is the Ricci formula

$$\lambda_{h/rs} = \sum \gamma_{hij} \lambda_{i/r} \lambda_{j/s} ,$$

but such expansion is not needed in the vector forms.

21. Let

$$\Theta = \tfrac{1}{2} \sum A_2 \cdot \alpha_i \phi_{\alpha_i}(\,)$$

be an operator which converts vectors of grade unity into vectors of grade 2. Of course it is a very special operator of this kind. The coefficient of rotation

$$\gamma_{ijk} = A_0 \cdot \alpha_i \alpha_j A_2 \cdot \Theta \alpha_k .$$

We see that $\Theta d\rho = \frac{1}{2} \sum A_2 \cdot \alpha_i \, d\alpha_i$ gives what may be called the instantaneous rotation for a displacement $d\rho$ about a plane. That is to say, the plexus of congruences rotates about this plane through an infinitesimal amount given by the intensity of the 2-vector. (Rotation about a plane or other spread is possible in hyperspace.) We can look at the situation in this way: let

$$\tau = A_1 \cdot \sigma \Theta \eta = - \sum \phi_{\alpha_j} \eta A_0 \cdot \alpha_j \sigma$$

and $A_0 \cdot \alpha_k \tau = A_0 \cdot \sigma \phi_{\alpha_k} \eta = A_0 \cdot \sigma d\alpha_k / dt$, $A_0 \cdot \sigma\tau = 0$. The coefficients of rotation are the components in the directions of the elementary bivectors $A_2 \cdot \alpha_i \alpha_j$ of the bivector $\Theta\alpha_k$.

We can also express Θ a little differently in a form useful in the study of n-tuply orthogonal congruences.*

Let

$$\zeta = \sum A_2 \cdot \nabla' \alpha_i' A_0 \cdot \alpha_i (\,) = - \sum A_2 \cdot \nabla' \alpha_i A_0 \cdot \alpha_i' (\,) ,$$

the accents showing the operand of ∇. Then

$$\Theta - \zeta = \tfrac{1}{2} A_2 (\,) \sum A_3 \cdot \alpha_i \alpha_i' \nabla' = - \tfrac{1}{2} A_2 \cdot (\,) \sum A_3 \, \alpha_i \nabla \alpha_i .$$

The expression $A_2 \cdot \nabla\alpha_i$ is of course the curl of α_i.

22. If $f(\rho) = 0$ is a spread of order $n-1$, the normals are ∇f and form a normal congruence. Since $f_{hk} = f_{kh}$, we have

$$A_0 \cdot \alpha_h \nabla A_0 \cdot \alpha_k \nabla \cdot f - A_0 \cdot \alpha_k \nabla A_0 \cdot \alpha_h \nabla \cdot f = 0 .$$

* See Shaw, *Triply orthogonal congruences*, these Transactions, vol. 21 (1920), p. 391. The present articles contain the results of a paper read December 31, 1915, before the Chicago section, *On orthogonal congruences in space of three and more dimensions.*

As ∇ operates upon following α's we have from this

$$A_0 \cdot \nabla f (\phi_{\alpha_i} \alpha_k - \phi_{\alpha_i} \alpha_h) = 0 .$$

This might be written

$$\sum (\gamma_{ikh} - \gamma_{ihk}) A_0 \cdot \alpha_i \nabla f = 0 .$$

Let the unit part of ∇f be α_n; then the equation gives

$$A_0 \cdot (\alpha_n \phi_{\alpha_n} \alpha_k - \alpha_n \phi_{\alpha_n} \alpha_h) = 0 .$$

That is, for the entire set of curves in the spread and normal to it,

$$\gamma_{nkh} = \gamma_{nhk} \qquad (h , k = 1 , 2 , \cdots , n - 1) .$$

The equation also states that ϕ_{α_i} is self-transverse (symmetric) so far as the spread is concerned. The condition is also sufficient. For when it is satisfied $A_0 \cdot A_2 \cdot \alpha_h \alpha_k A_2 \cdot \nabla \alpha_0 = 0$ for all $h , k = 1 , \cdots , n - 1$. But

$$A_2 \nabla \alpha_n = \sum x_j A_2 \alpha_n \alpha_j ,$$

hence we have

$$A_3 \cdot \alpha_n \nabla \alpha_n = 0 ,$$

whence $t A_2 \cdot \nabla \alpha_n = A_2 \xi \alpha_n$ for any ξ. Choose $\xi = - \nabla t$ and we have

$$A_2 \nabla (t \alpha_n) = 0 , \qquad t \alpha_n = \nabla f .$$

From this we have also $A_2 \cdot \nabla \log t \alpha_n = - A_2 \cdot \nabla \alpha_n = A_2 \cdot \alpha_n \pi$, where π is in the spread and is indeed $\phi_{\alpha_n} \alpha_n$.

23. Let us suppose now that α_n is the principal normal of a system of curves through the point, that is, α_i is the unit tangent, $d\alpha_i / ds$ for $d\rho$ along the curve is $c_i \alpha_n$, where $i = 1 , \cdots , n - 1$. That is,

$$\phi_{\alpha_i} \alpha_i = c_i \alpha_n \quad \text{for all} \quad i = 1 , \cdots , n - 1 .$$

Hence

$$A_2 \cdot \alpha_n \phi_{\alpha_i} \alpha_i = 0 , \qquad A_0 \cdot \alpha_j \phi_{\alpha_i} \alpha_i = - A_0 \cdot \alpha_i \phi_{\alpha_j} \alpha_i = 0 \qquad (i , j = 1 , \cdots , n - 1) .$$

This may also be written to parallel the treatment in three variables referred to above:

$$A_1 \cdot \alpha_i \Theta \alpha_i = c_i \alpha_n ,$$
$$A_0 \cdot \alpha_n \phi_{\alpha_i} \alpha_i = c_i .$$

If the curves all lie in a spread as in the preceding article,

$$\nabla \log t = \phi_{\alpha_n} \alpha_n \tau + x \alpha_n ,$$
$$A_0 \cdot \alpha_n \phi_{\alpha_i} \alpha_j = A_0 \cdot \alpha_n \phi_{\alpha_j} \alpha_i .$$

In the general case of all the curves merely such that α_n is their common principal normal, the sum of the curvatures is

$$A_0 \cdot \alpha_n \sum \phi_{\alpha_i} \alpha_i \qquad (i = 1 , \cdots , n) .$$

which reduces to

$$-\sum A_0 \cdot \alpha_n A_1 \cdot \alpha_i \Theta\alpha_i \qquad \text{or} \qquad \sum - A_0 \cdot \alpha_i \phi_{\alpha_n} \alpha_i = - A_0 \cdot \nabla\alpha_0 .$$

This shows on the one hand that the sum of the curvatures of the normal curves is independent of the particular system chosen, if they are mutually orthogonal, but on the other hand it gives significance to the geometric convergence of α_n. Dividing by $n-1$ would give the mean curvature of the system.

The second curvature of the curve with α_i as tangent is

$$\phi_{\alpha_n} \alpha_i = c_i'' \gamma_i - c_i' \alpha_i .$$

Since α and γ are orthogonal,

$$A_2 \cdot \alpha_i \phi_{\alpha_n} \alpha_i = c_i'' \alpha_i \gamma_i ,$$

and we may write

c'' is the intensity of $A_2 \cdot \alpha_i \phi_{\alpha_n} \alpha_i$.

24. The spread of order $n-1$ is called *isothermal* when

$$A_0 \nabla\nabla f = 0 = \nabla^2 f .$$

The resulting equations are simple.

25. In case the curves in the region through the point in question are geodesic we have $\phi_{\alpha_n} \alpha_n = 0$ in the region, that is, for every i

$$A_0 \cdot \alpha_i \phi_{\alpha_n} \alpha_n = 0 .$$

That is,

$$A_0 \cdot \alpha_n \phi_{\alpha_i} \alpha_n = 0$$

for every i. This may also be written $A_0 \cdot \alpha_n \alpha_i \Theta\alpha_n = 0$. This enables us to state the condition in the form $A_1 \cdot \alpha_n \Theta\alpha_n = 0$ in the region, and this vector must lie outside the region therefore.*

Further applications must be deferred.

*See page 213.

UNIVERSITY OF ILLINOIS,
URBANA, ILL.

REPRESENTATIONS OF A COMPLEX POINT BY PAIRS OF ORDERED REAL POINTS *

BY

W. C. GRAUSTEIN

1. **Introduction.** A complex point z in a complex space of n dimensions, since it depends on $2n$ real parameters, can be represented by a pair of real points, X and Y, of the space. It is reasonable to restrict the choice of X and Y by the following conditions:

A. The points X and Y representing z shall be ordered: $X \rightarrow Y$.

B. If z is a real point, X and Y shall coincide in z.

C. If z is imaginary, X and Y shall be distinct; if $X \rightarrow Y$ represents z, $Y \rightarrow X$ shall represent $\bar{z}$, the point conjugate-imaginary to z.

D. The coördinates of X and Y, referred to a rectangular coördinate system, shall be analytic functions, without singularities in the finite domain, of the coördinates of z and $\bar{z}$, referred to this system.

E. The representation of z by $X \rightarrow Y$ shall be invariant under a chosen group of real point transformations; that is, the ordered points $X' \rightarrow Y'$ into which the ordered points $X \rightarrow Y$ are carried by an arbitrary transformation of the group shall represent the point z' into which z is carried by the transformation.

If $X \rightarrow Y$ is a representation of z which satisfies these conditions, $Y \rightarrow X$ is also a representation of z satisfying them. The two representations shall be called *inverse*.

There are two pairs of inverse representations satisfying the prescribed conditions which are well known: those of Laguerre in the plane, whereby z is represented by the real points of intersection, taken in the one or the other order, of the minimal lines through z and $\bar{z}$; and those of Marie in the plane and in space, in which the real points of the representation lie on the line $z\bar{z}$, have the same midpoint as z and $\bar{z}$, and are at a distance apart whose square is the negative of the square of the distance $z\bar{z}$.† The two Marie representations are invariant under the group of affine transformations, whereas those of Laguerre are unchanged by the direct circular transformations of the plane.

This paper deals with the determination of all the representations $X \rightarrow Y$

* Presented to the Society, December 31, 1919.

† For a review of the literature concerning these representations, see E. Study, *Vorlesungen über ausgewählte Gegenstände der Geometrie, erstes Heft*; *Ebene analytische Kurven und zu ihnen gehörige Abbildungen*, Leipzig, 1911; J. L. Coolidge, *Geometry in the Complex Domain* (in press).

of z which satisfy conditions A to D, and are invariant, in compliance with E, under a given group. In § 2 the problem is formulated analytically for an arbitrary group. In the subsequent sections, 3–7, it is solved for various subgroups of the projective group, for the projective group itself, and for the group of direct circular transformations of the plane. Only in the case of the projective group are no representations obtained.

Given a group, G, of direct transformations and the set, S, of corresponding indirect transformations, it is of importance to determine, among the representations invariant under G, those which are invariant also under S, and secondly, those which are inverted, i.e., carried into their inverses, by S. It turns out that all the representations invariant under G are invariant also under S, except when G is the group of direct circular transformations, or when G is either the group of motions or of direct transformations of similarity in the plane.* In the latter cases certain representations of the totality invariant under G are unchanged by S, whereas certain others are inverted by S (§ 8). Moreover, it is only in these same cases that the representations are of exceptional nature in comparison with those obtained for the same group applied in a space of different dimensionality.

2. **Analytic formulation of the problem.** Let $(z_1, z_2, \cdots, z_n)$, where

$$z_i = x_i + iy_i \qquad (i = 1, 2, \cdots, n),$$

be the coördinates, referred to a rectangular cartesian system of axes, of the point z, and let the coördinates of X and Y, referred to the same system, be respectively $(X_1, X_2, \cdots, X_n)$ and $(Y_1, Y_2, \cdots, Y_n)$.

Since the latter sets of coördinates are, by condition D, real analytic functions of $z_1, z_2, \cdots, z_n, \bar{z}_1, \bar{z}_2, \cdots, \bar{z}_n$, they are also real analytic functions of the real variables $x_1, y_1, x_2, y_2, \cdots, x_n, y_n$. It is convenient, in light of B, to write these functions as follows:

$$X_i = x_i + F_i(x_1, y_1, \cdots, x_n, y_n),$$
$$Y_i = x_i + G_i(x_1, y_1, \cdots, x_n, y_n) \qquad (i = 1, 2, \cdots, n).$$

By C,

$$G_i(x_1, y_1, \cdots, x_n, y_n) = F_i(x_1, -y_1, \cdots, x_n, -y_n) \qquad (i = 1, 2, \cdots, n).$$

Conditions A to D are thus covered by demanding that X_i and Y_i be of the form

$$\begin{aligned} X_i &= x_i + F_i(x_1, y_1, \cdots, x_n, y_n), \\ Y_i &= x_i + F_i(x_1, -y_1, \cdots, x_n, -y_n) \qquad (i = 1, 2, \cdots, n), \end{aligned} \tag{1}$$

where $F_1, F_2, \cdots, F_n$ are real functions of the real variables $x_1, y_1, \cdots, x_n, y_n$, which are analytic for all finite sets of values, vanish identically when

$$y_1 = y_2 = \cdots = y_n = 0,$$

* In this summary of exceptions, one relatively unimportant case has been omitted.

and are furthermore such that the system of equations

$$F_i(x_1, y_1, \cdots, x_n, y_n) = F_i(x_1, -y_1, \cdots, x_n, -y_n)$$

has no real solutions in $y_1, y_2, \cdots, y_n$ except the solution $0, 0, \cdots, 0$.

Let the equations

$$z_i' = T_i(z_1, z_2, \cdots, z_n) \qquad (i = 1, 2, \cdots, n) \tag{2}$$

represent an arbitrary group of real point transformations. In order that the representation $X \to Y$ of the point z be invariant under this group, it is necessary and sufficient that

$$\begin{aligned} &(a)\ x_i' + F_i(x_1', y_1', \cdots, x_n', y_n') \\ &\quad = T_i(x_1 + F_1(x_1, y_1, \cdots, x_n, y_n), \cdots, \\ &\qquad x_n + F_n(x_1, y_1, \cdots, x_n, y_n)), \\ &(b)\ x_i' + F_i(x_1', -y_1', \cdots, x_n', -y_n') \\ &\quad = T_i(x_1 + F_1(x_1, -y_1, \cdots, x_n, -y_n), \cdots, \\ &\qquad x_n + F_n(x_1, -y_1, \cdots, x_n, -y_n)) \quad (i = 1, 2, \cdots, n). \end{aligned} \tag{3}$$

Thus our problem, finally formulated, consists in solving the system of functional equations (3) for the functions $F_1, F_2, \cdots, F_n$, subject to the conditions stated under (1).

If the transformations (2) are linear, as will frequently be the case,

$$x_i' = T_i(x_1, x_2, \cdots, x_n), \qquad y_i' = T_i^0(y_1, y_2, \cdots, y_n) \qquad (i = 1, 2, \cdots, n),$$

where T_i^0 is T_i minus its absolute term. Changing the signs of $y_1, y_2, \cdots, y_n$ then changes the signs of $y_1', y_2', \cdots, y_n'$. Consequently, the two systems of equations in (3) are, in this case, each equivalent to the single system

$$\begin{aligned} F_i(x_1', y_1', \cdots, x_n', y_n') = T_i^0(&F_1(x_1, y_1, \cdots, x_n, y_n), \\ &\cdots, F_n(x_1, y_1, \cdots, x_n, y_n)) \quad (i = 1, 2, \cdots, n). \end{aligned} \tag{4}$$

3. **Representations invariant under the group of translations.** Inasmuch as the group of real translations is a subgroup of each of the other groups under which the invariance of representations is to be considered, it is advantageous to deal with it first. Equations (4) then become

$$F_i(x_1 + a_1, y_1, \cdots, x_n + a_n, y_n) = F_i(x_1, y_1, \cdots, x_n, y_n) \qquad (i = 1, 2, \cdots, n).$$

Since $a_1, a_2, \cdots, a_n$, the parameters of the translation, are arbitrary, the system is satisfied if and only if $F_1, F_2, \cdots, F_n$ are functions of $y_1, y_2, \cdots, y_n$ alone.

The representations $X \to Y$ of z invariant under the group of real translations

are given by

$$\text{I} \qquad X_i = x_i + F_i(y_1, y_2, \cdots, y_n), \quad Y_i = x_i + F_i(-y_1, -y_2, \cdots, -y_n) \quad (i = 1, 2, \cdots, n),$$

where $F_1, F_2, \cdots, F_n$ are arbitrary real analytic functions which vanish for $y_1 = y_2 = \cdots = y_n = 0$ and are such that the system of equations

$$F_i(y_1, y_2, \cdots, y_n) = F_i(-y_1, -y_2, \cdots, -y_n) \quad (i = 1, 2, \cdots, n)$$

has no real solutions other than $0, 0, \cdots, 0$.

In the case of a group of linear transformations containing the group of translations, (4) can now be replaced by the system of equations

$$(5) \qquad \begin{aligned} F_i(y_1', y_2', \cdots, y_n') = T_i^0(&F_1(y_1, y_2, \cdots, y_n), \\ &\cdots, F_n(y_1, y_2, \cdots, y_n)) \quad (i = 1, 2, \cdots, n), \end{aligned}$$

in which $x_1, x_2, \cdots, x_n, x_1', x_2', \cdots, x_n'$ no longer appear.*

4. **The group of homothetic transformations.** The representation $X \to Y$ is invariant under the group of real homothetic transformations

$$z_i' = \rho z_i + a_i, \qquad \rho \neq 0 \qquad (i = 1, 2, \cdots, n),$$

if and only if the system (5), formed for this group, namely

$$F_i(\rho y_1, \rho y_2, \cdots, \rho y_n) = \rho F_i(y_1, y_2, \cdots, y_n) \quad (i = 1, 2, \cdots, n),$$

is satisfied. But then, regardless of whether ρ is unrestricted or takes on only positive values, $F_1, F_2, \cdots, F_n$ must be linear and homogeneous in $y_1, y_2, \cdots, y_n$.

The representations $X \to Y$ of z invariant under the group of either direct or general homothetic transformations are

$$\text{II} \qquad X_i = x_i + \sum_{j=1}^{n} k_{ij} y_j, \qquad Y_i = x_i - \sum_{j=1}^{n} k_{ij} y_j \quad (i = 1, 2, \cdots, n),$$

where the k_{ij} are real constants subject only to the restriction that the n-rowed determinant $|k_{ij}|$ does not vanish.

* The equations (5) are also necessary and sufficient conditions that the real analytic point transformation

$$(a) \qquad y_i' = F_i(y_1, y_2, \cdots, y_n) \qquad (i = 1, 2, \cdots, n)$$

be commutative with each transformation of the group of real linear transformations

$$(b) \qquad y_i' = T_i^0(y_1, y_2, \cdots, y_n) \qquad (i = 1, 2, \cdots, n)$$

leaving the origin invariant. Accordingly, as we solve (5) in the subsequent paragraphs for various groups (b), we determine at the same time, for each particular group (b), the group of transformations (a), each of which is commutative with every transformation (b). Thus we find, for example, in § 6, that the group of transformations commutative with the group of motions about the origin in the plane is

$$y_1' = y_1 R_1(r^2) + y_2 R_2(r^2), \qquad y_2' = -y_1 R_2(r^2) + y_2 R_1(r^2),$$

where $r^2 = y_1^2 + y_2^2$ and R_1, R_2 are analytic functions of r^2, not both zero.

The condition $|k_{ij}| \neq 0$ is necessary and sufficient, not merely that X and Y coincide only when z is real, but also that the $2n$ equations II can be solved uniquely for $x_1, x_2, \cdots, x_n, y_1, y_2, \cdots, y_n$ and hence the coördinates of z found in terms of those of X and Y. In other words, the correspondence between the ordered point pairs $X \rightarrow Y$ and the points z under a representation II is one-to-one without exception in the finite domain.

5. **The affine and projective groups.** In order that the representation II be invariant under the n transformations

$$z_q' = -z_q, \qquad z_i' = z_i \qquad (q = 1, 2, \cdots, n),$$

where in each transformation i takes on all the integral values from 1 to n except the fixed value q, it is necessary that $k_{ij} = 0$, $i \neq j$. If the resulting representation is to be unchanged by each of the transformations

$$z_q' = z_r, \qquad z_r' = z_q, \qquad z_i' = z_i \qquad (q \neq r),$$

where in each transformation q and r are fixed and i takes on the remaining values from 1 to n, it follows that $k_{11} = k_{22} = \cdots = k_{nn}$. A necessary condition, therefore, that II be invariant under the affine group is that $k_{ij} = 0$, $i \neq j$, and $k_{ii} = k$ for all i. That this condition is also sufficient can be easily verified.

If, instead of all real affine transformations, merely those for which the determinant of the coefficients is positive had been admitted, the result would have been the same.

The representations $X \rightarrow Y$ of z invariant under the group of either direct or general affine transformations are

III $$X_i = x_i + ky_i, \qquad Y_i = x_i - ky_i \qquad (i = 1, 2, \cdots, n),$$

where k is an arbitrary constant, $\neq 0$.

The two inverse representations for which $k = \pm 1$ are *the representations of Marie.*

If III is to be invariant under the projective group, it must be unchanged by the transformation

$$z_1' = \frac{1}{z_1}, \qquad z_i' = \frac{z_i}{z_1} \qquad (i = 2, \cdots, n).$$

But then the first of the equations (3) necessitates that $k^2 = -1$.

There is no representation $X \rightarrow Y$ of z which is invariant under the projective group.

6. **The group of motions.** The treatment in this case varies with the dimensionality of the space.

A. $n = 1$. The group of motions is the group of translations. According to I, *the invariant representations are*

IVA $$X_1 = x_1 + y_1 \Phi(y_1), \qquad Y_1 = x_1 - y_1 \Phi(y_1),$$

where $\Phi(y_1)$ is a real analytic function and the equation $\Phi(y_1) = -\Phi(-y_1)$ is satisfied by no real value of y_1 except perhaps $y_1 = 0$.

B. $n = 2$. The representation must be of the form

$$X_i = x_i + F_i(y_1, y_2), \qquad Y_i = x_i + F_i(-y_1, -y_2) \qquad (i = 1, 2), \tag{6}$$

and be invariant under the transformations

$$z_1' = \cos\phi\, z_1 - \sin\phi\, z_2 + a_1, \qquad z_2' = \sin\phi\, z_1 + \cos\phi\, z_2 + a_2.$$

That is, $F_1(y_1, y_2)$, $F_2(y_1, y_2)$ must be solutions of the system of equations, obtained from (5),

$$\begin{aligned} F_1(y_1', y_2') &= \cos\phi\, F_1(y_1, y_2) - \sin\phi\, F_2(y_1, y_2), \\ F_2(y_1', y_2') &= \sin\phi\, F_1(y_1, y_2) + \cos\phi\, F_2(y_1, y_2), \end{aligned} \tag{7}$$

where

$$y_1' = \cos\phi\, y_1 - \sin\phi\, y_2, \qquad y_2' = \sin\phi\, y_1 + \cos\phi\, y_2.$$

For $\phi = \pi$, equations (7) become the conditions that F_1, F_2 be odd functions of y_1, y_2. Consequently, if F_i is written in the form

$$F_i(y_1, y_2) = \frac{F_i(y_1, y_2) + F_i(y_1, -y_2)}{2} + \frac{F_i(y_1, y_2) + F_i(-y_1, y_2)}{2} \qquad (i = 1, 2),$$

the first quotient on the right is even in y_2, odd in y_1 and hence identically zero for $y_1 = 0$, whereas the second is even in y_1, odd in y_2 and so identically zero for $y_2 = 0$. Accordingly,

$$F_i(y_1, y_2) = y_1 F_i'(y_1, y_2) + y_2 F_i''(y_1, y_2) \qquad (i = 1, 2), \tag{8}$$

where each of the functions F_1', F_1'', F_2', F_2'' is analytic and is even in y_1, in y_2, and hence in y_1, y_2.

In the equations obtained from (7) by substituting for F_1 and F_2 their values as given by (8), replace ϕ by $-\phi$, y_1 by $-y_1$, and combine by addition and subtraction each of the resulting equations with that from which it arose. The system obtained is

$$\begin{aligned} y_1' F_1'(y_1', y_2') &= y_1 \cos\phi\, F_1'(y_1, y_2) - y_2 \sin\phi\, F_2''(y_1, y_2), \\ y_2' F_2''(y_1', y_2') &= y_1 \sin\phi\, F_1'(y_1, y_2) + y_2 \cos\phi\, F_2''(y_1, y_2), \\ y_1' F_2'(y_1', y_2') &= y_1 \cos\phi\, F_2'(y_1, y_2) + y_2 \sin\phi\, F_1''(y_1, y_2), \\ y_2' F_1''(y_1', y_2') &= -y_1 \sin\phi\, F_2'(y_1, y_2) + y_2 \cos\phi\, F_1''(y_1, y_2). \end{aligned} \tag{9}$$

These equations reduce, for $\phi = \pi/2$, to

$$F_2'(y_1, y_2) = -F_1''(y_2, y_1), \qquad F_2''(y_1, y_2) = F_1'(y_2, y_1). \tag{10}$$

When the values for F_2' and F_2'' given by (10) are substituted in (9) and y_2 is

set equal to zero, the four resulting equations yield immediately the identities

$$F_1'(y_1 \cos\phi, y_1 \sin\phi) = F_1'(y_1 \sin\phi, y_1 \cos\phi),$$
$$F_1''(y_1 \cos\phi, y_1 \sin\phi) = F_1''(y_1 \sin\phi, y_1 \cos\phi),$$

in the independent variables y_1 and ϕ. Consequently, $F_1'(y_1, y_2)$ and $F_1''(y_1, y_2)$ are symmetric functions of y_1, y_2, and (10) becomes

$$F_2'(y_1, y_2) = -F_1''(y_1, y_2), \qquad F_2''(y_1, y_2) = F_1'(y_1, y_2). \tag{11}$$

But the system (9) then reduces to

$$F_1'(y_1', y_2') = F_1'(y_1, y_2), \qquad F_1''(y_1', y_2') = F_1''(y_1, y_2).$$

That is, F_1' and F_1'' are absolute invariants of the group of rotations about the origin, and hence must be of the form

$$F_1' = R_1(r^2), \qquad F_1'' = R_2(r^2), \tag{12}$$

where

$$r^2 = y_1^2 + y_2^2.$$

The necessary conditions (S), (11), (12) that $F_1(y_1, y_2)$, $F_2(y_3, y_2)$ be solutions of (7) are shown sufficient by direct substitution.

The representations $X \rightarrow Y$ of z invariant under the group of motions in the plane are

$$\text{IVB} \quad \begin{aligned} X_1 &= x_1 + y_1 R_1(r^2) + y_2 R_2(r^2), \quad Y_1 = x_1 - y_1 R_1(r^2) - y_2 R_2(r^2), \\ X_2 &= x_2 - y_1 R_2(r^2) + y_2 R_1(r^2), \quad Y_2 = x_2 + y_1 R_2(r^2) - y_2 R_1(r^2), \end{aligned}$$

where $r^2 = y_1^2 + y_2^2$, and R_1, R_2 are real functions of r^2, analytic for all finite values and not vanishing simultaneously except perhaps when $r^2 = 0$.

C. $n \geqq 3$. In case $n = 3$, the functions F_1, F_2, F_3 in the representation

$$X_i = x_i + F_i(y_1, y_2, y_3), \qquad Y_i = x_i + F_i(-y_1, -y_2, -y_3) \qquad (i = 1, 2, 3) \tag{13}$$

are to be determined from the system of equations

$$F_i(y_1', y_2', y_3') = \sum_{j=1}^{3} a_{ij} F_j(y_1, y_2, y_3) \qquad (i = 1, 2, 3), \tag{14}$$

where

$$y_i' = \sum_{j=1}^{3} a_{ij} y_j \qquad (i = 1, 2, 3),$$

and the determinant $|a_{ij}|$ is orthogonal and has the value $+1$.

In case of an arbitrary rotation about the axis of z_3, $y_3' = y_3$ and y_1, y_2 are transformed as for $n = 2$. The last equation of (14) reduces to

$$F_3(y_1', y_2', y_3) = F_3(y_1, y_2, y_3),$$

and the first two to the equations (7), with the variable y_3 added in each

function. Consequently,

$$
\begin{aligned}
F_1 &= y_1 F_3'(y_1^2 + y_2^2, y_3) + y_2 F_3''(y_1^2 + y_2^2, y_3), \\
F_2 &= -y_1 F_3''(y_1^2 + y_2^2, y_3) + y_2 F_3'(y_1^2 + y_2^2, y_3), \\
F_3 &= F_3'''(y_1^2 + y_2^2, y_3).
\end{aligned} \tag{15}
$$

Corresponding respectively to the rotations about the axes of z_1 and z_2, there are two other sets of equations like (15) and obtainable from (15) by permuting the subscripts 1, 2, 3 cyclicly. From the three sets of equations it follows that $F_1'' = F_2'' = F_3'' = 0$; for example, since $F_1 = F_1'''(y_2^2 + y_3^2, y_1)$, F_1 is an even function of y_2, whence, by the first equation of (15), $F_2'' = 0$. Therefore

$$F_i = y_i F_j'(y_i^2 + y_k^2, y_j) = y_i F_k'(y_i^2 + y_j^2, y_k),$$

where i, j, k are to be permuted cyclicly through the values 1, 2, 3. Thus

$$F_1' = F_2' = F_3' = R(r^2), \qquad \text{where} \qquad r^2 = y_1^2 + y_2^2 + y_3^2,$$

and (13) becomes

$$X_i = x_i + y_i R(r^2), \qquad Y_i = x_i - y_i R(r^2) \qquad (i = 1, 2, 3).$$

Substitution shows that the functions F_i in this representation satisfy (14).

This result for $n = 3$ can be generalized by mathematical induction to hold for the case $n > 3$.

The representations $X \rightarrow Y$ of z invariant under the group of motions in n-dimensional space, where $n \geqq 3$, are

IVC $$X_i = x_i + y_i R(r^2), \qquad Y_i = x_i - y_i R(r^2) \quad (i = 1, 2, \cdots, n),$$

where

$$r^2 = y_1^2 + y_2^2 + \cdots + y_n^2,$$

and $R(r^2)$ is a real function analytic for all finite values and not zero except perhaps when $r^2 = 0$.

For the group of motions and reflections the result incorporated in IVC *holds for all values of n.* To establish the truth of this statement for $n \geqq 3$, it is sufficient to show that the representation IVC is invariant under a reflection in one of the coördinate hyperplanes. If $n = 2$, the representation must be of the form IVB and be invariant under $z_1' = z_1$, $z_2' = -z_2$; then $R_2(r^2) = 0$, and conversely. Finally, IVA is invariant under $z_1' = -z_1$ if and only if $\Phi(y_1)$ is an even function of y_1, i.e., is of the form $R(y_1^2)$.

7. **The group of direct transformations of similarity.** Necessary and sufficient that the representation $X \rightarrow Y$ be unchanged by the direct transformations of similarity in the cases $n = 1$, $n = 2$, $n \geqq 3$ is the requirement that IVA, IVB, IVC be respectively invariant under an arbitrary stretching from the origin of positive ratio. According to § 4, the arbitrary functions involved in the representations IV are, then, all constants, and conversely.

The representations $X \rightarrow Y$ of z invariant under the group of direct transformations of similarity are

$$\text{VA} \qquad X_1 = x_1 + ky_1, \qquad Y_1 = x_1 - ky_1, \qquad k \neq 0, \quad \text{if} \quad n = 1,$$

$$\text{VB} \qquad \begin{aligned} X_1 &= x_1 + ky_1 + ly_2, & Y_1 &= x_1 - ky_1 - ly_2, \\ X_2 &= x_2 - ly_1 + ky_2, & Y_2 &= x_2 + ly_1 - ky_2, \end{aligned} \qquad k^2 + l^2 \neq 0, \quad \text{if} \quad n = 2,$$

$$\text{VC} \qquad X_i = x_i + ky_i, \qquad Y_i = x_i - ky_i, \qquad k \neq 0, \quad i = 1, 2, \cdots, n, \quad \text{if} \quad n \geqq 3.$$

Except in the plane, these representations are identical with the representations III invariant under the affine group.

For the group of direct and indirect transformations of similarity the representations are for all n given by VC *and are identical with those peculiar to the affine group.*

The two inverse representations VB, for which $k = 0$ and $l = \pm 1$, namely

$$\text{VI} \qquad \begin{aligned} X_1 &= x_1 \pm y_2, & Y_1 &= x_1 \mp y_2, \\ X_2 &= x_2 \mp y_1, & Y_2 &= x_2 \pm y_1, \end{aligned}$$

are *the representations of Laguerre*. From their geometrical definition (§ 1), it is evident that these representations are invariant under the direct circular transformations. It can be shown that they are the only representations with this property.

8. **Representations inverted by a set of transformations.** Given a group, G, of direct transformations of a certain type and the set, S, of indirect transformations of the same type, to determine, among the representations invariant under G, those which, though not invariant under S, are merely inverted by the transformations of S.

It is geometrically evident that a Laguerre representation is inverted by an indirect circular transformation. Of the groups G other than the circular group, every one under which invariant representations exist is linear and contains the group of translations. Hence the representation must be of the form I. In order that it be inverted by a linear transformation (2), it is necessary and sufficient that the functions F_i satisfy the system of equations

$$\begin{aligned} &F_i(-y_1', -y_2', \cdots, -y_n') \\ &\quad = T_i^0(F_1(y_1, y_2, \cdots, y_n), \cdots, F_n(y_1, y_2, \cdots, y_n)) \\ &\qquad (i = 1, 2, \cdots, n), \end{aligned} \tag{16}$$

In §§ 4–7 it has been shown that, when G is any one of the following groups of direct transformations: homothetic group, affine group, group of motions

($n \geqq 3$), similarity group ($n \neq 2$), all the representations invariant under G are also invariant under S. It is true also that no one of the representations IVA, invariant under the group of motions for $n = 1$, is inverted by any transformation of S, as can readily be proved by applying (16) with $z_1' = -z_1$ as the transformation used; but in this case, not all the representations are invariant under S (§ 6).

If G is the group of motions (or of direct transformations of similarity) in the plane, the functions F_i to be used in (16) are as in IVB (or VB). It is sufficient to choose as the transformation a reflection in a coördinate axis. Then equations (16) reduce to $R_1 = 0$ (or $k = 0$).

Of the representations IVB *invariant under the group of motions in the plane, only those for which* $R_2 = 0$ *are invariant under reflections, whereas only those for which* $R_1 = 0$ *are inverted by every reflection.*

Of the representations VB *invariant under the direct transformations of similarity, only those for which* $l = 0$ *are invariant under the indirect transformations of similarity, whereas only those for which* $k = 0$ *are inverted by every such transformation.*

HARVARD UNIVERSITY,
CAMBRIDGE, MASS.

NON-LOXODROMIC SUBSTITUTIONS AND GROUPS IN N DIMENSIONS*

BY

EDWARD B. VAN VLECK

Linear substitutions

$$x' = \frac{ax + b}{cx + d} \tag{1}$$

are commonly divided into two classes, loxodromic and non-loxodromic, and the non-loxodromic substitutions are further classified as elliptic, parabolic, and hyperbolic. So far as I am aware, no corresponding classification of projective substitutions in $n - 1$ nonhomogeneous or n homogeneous variables has been suggested.† The classification which is introduced below is based upon the formulation of a projective substitution under the shape of a homogeneous substitution with determinant $+1$. If the substitution (1) is written in homogeneous form

$$x_1' = ax_1 + bx_2, \qquad x_2' = cx_1 + dx_2$$

with unit determinant, the substitution is non-loxodromic when, and only when, $a + d$ is real; that is, when the characteristic equation

$$\begin{vmatrix} a - \lambda & b \\ c & d - \lambda \end{vmatrix} = 0$$

is a *real* equation. The substitution is then hyperbolic, parabolic, or elliptic ‡ according as the two roots of this equation are real and distinct, coincident, or imaginary. Correspondingly, the projective substitutions are taken *exclusively* in this paper in homogeneous form,

$$\begin{aligned} x_1' &= a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n, \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ x_n' &= a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n, \end{aligned} \qquad (S) \tag{2}$$

* Presented to the Society, March 25, 1921.

† Fubini in his *Introduzione alla Teoria dei Gruppi discontinui e delle Funzioni automorfe*, pp. 77–8, has applied the above terms to those substitutions which are movements of

$$x_n^2 + \epsilon(x_1^2 + x_2^2 + \cdots + x_{n-1}^2) = 1 \qquad (\epsilon = \pm 1),$$

into itself. His aims and distinctions are very different from those of this paper.

‡ An application of the terms elliptic and hyperbolic differing slightly from the above usage of Klein is also to be found in the literature. Cf. Veblen's *Projective Geometry*, vol. 2, pp. 5, 171, 248.

with determinant Δ equal to $+1$. I call a substitution non-loxodromic, if, when so written, it possesses a real characteristic equation,

$$\begin{vmatrix} a_{11}-\lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22}-\lambda & \cdots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \cdots & a_{nn}-\lambda \end{vmatrix} = 0. \tag{3}$$

The characteristic roots are accordingly real or pairs of conjugate imaginaries.

It is well known that there are just two classes of groups (1) in a single variable which consist of non-loxodromic substitutions without a common pole. Either all the substitutions of the group leave invariant a circle in the complex x-plane and the group is called a "group with a principal circle" (Hauptkreis Gruppe), or the group consists entirely of elliptic substitutions. In the former case the group can be so transformed that the circle becomes the real axis. We have then a canonical form for the group in which its substitutions have real coefficients and unit-determinant, and, conversely, any group of this character is non-loxodromic. A second canonical form is

$$x' = \frac{(a+bi)x + (c+di)}{(c-di)x + (a-bi)} \qquad (a^2+b^2-c^2-d^2=1), \tag{3'}$$

which leaves the unit circle invariant. The second class of non-loxodromic groups—the elliptic groups—can be put into a canonical form in which the substitutions are all of form

$$x' = \frac{(a+bi)x + (c+di)}{-(c-di)x + (a-bi)} \qquad (a^2+b^2+c^2+d^2=1), \tag{4}$$

and the group can be interpreted as a group of rotations of a sphere into itself.

When the number of homogeneous variables is greater than 2, there are, of course, more varieties of non-loxodromic substitutions to be distinguished than the familiar hyperbolic, elliptic and parabolic types, and it was perhaps scarcely to be expected that the result summarized in the preceding paragraph would have its exact analogue. Nevertheless an exactly similar result is obtained. There are still just two types of non-loxodromic groups. A limitation upon this statement (perhaps unnecessary) is introduced by the proof which requires that the non-loxodromic group should contain at least one substitution having a characteristic equation with n distinct roots. Corresponding to the hypothesis of the preceding paragraph that the substitutions (1) should not have a common pole, we have now the requirement that our non-loxodromic group of substitutions (2) in n homogeneous variables should not leave invariant a linear space of less than $n-1$ (complex) dimensions; in other words, should not transform into itself a hyperplane nor a space common to two or more hyperplanes in our projective space S_{n-1}.

One normal form obtained for the first class of non-loxodromic groups is that in which all substitutions (2) of the group have real coefficients and a determinant equal to $+1$. In this normal form the group obviously leaves invariant the real space embedded in the complex space S_{n-1} of the n homogeneous complex coördinates which is of $2n-2$ real dimensions. Any projective $(n-1)$-dimensional image of this real space within the complex space has been called a "chain" (von Staudt, Pieri), and every non-loxodromic group of our first class accordingly leaves some chain invariant. If $n-1$ is even, all projective substitutions which leave a chain invariant form a non-loxodromic group of the first class. When $n-1$ is odd, this is not the case. The most inclusive non-loxodromic group which leaves the chain invariant is then a subgroup of the set of all substitutions (2) which leave the chain invariant.*

The second class of non-loxodromic groups is novel and interesting. It occurs only when n is an even number $2m$. Any group of this class can be so transformed that all its substitutions have simultaneously the canonical form

$$
\begin{aligned}
x'_{2i-1} &= (a_{2i-1,\,1}\,x_1 + a_{2i-1,\,2}\,x_2) \\
&\qquad + \cdots + (a_{2i-1,\,2m-1}\,x_{2m-1} + a_{2i-1,\,2m}\,x_{2m}), \\
x'_{2i} &= (-\bar{a}_{2i-1,\,2}\,x_1 + \bar{a}_{2i-1,\,1}\,x_2) \\
&\qquad + \cdots + (-\bar{a}_{2i-1,\,2m}\,x_{2m-1} + \bar{a}_{2i-1,\,2m-1}\,x_{2m}) \\
&\qquad\qquad (i = 1, \cdots, m),
\end{aligned} \tag{5}
$$

where $\bar{a}_{ij}$ denotes the conjugate of a_{ij}. Thus the determinant of the substitution consists of m^2 blocks of form

$$
\begin{vmatrix} a_{ij} & a_{i,\,j+1} \\ -\bar{a}_{i,\,j+1} & \bar{a}_{ij} \end{vmatrix} \qquad (\text{odd } i \text{ and } j), \tag{6}
$$

similar to the block of coefficients of (4). It is shown that when a determinant has this structure, it is necessarily positive or zero.†

For any (non-singular) substitution of the second class the elementary divisors of (3) occur only in conjugate pairs $(\lambda-\lambda_i)^s$, $(\lambda-\bar{\lambda}_i)^s$ of equal degree, this statement applying as well for real roots λ_i which furnish pairs of equal elementary divisors. Conversely, a non-singular substitution which fulfills this requirement is one of the second class and may be transformed into the normal form (5). Such substitutions will be called *elliptic (restrictedly* elliptic, when the characteristic roots are m distinct pairs of conjugate imag-

* J. W. Young in these Transactions, vol. 11 (1910), p. 280, investigates the necessary and sufficient condition that a collineation ($n=3$) shall leave a two-dimensional chain invariant.

† This result has been known for the special case in which all the elements of the determinant are real. Cf. Kowalewski, *Einführung in die Determinantentheorie*, pp. 156–158.

inaries). The characteristic property of an elliptic substitution in its normal form (5) is that it converts any pair of points

$$P = (x_1, x_2; \cdots; x_{2m-1}, x_{2m}), \qquad Q = (-\bar{x}_2, \bar{x}_1; \cdots; -\bar{x}_{2m}, \bar{x}_{2m-1})$$

into another pair of points whose coördinates are similarly related, any substitution having this property being necessarily of form (5). The invariant points of any substitution (5) occur in such pairs; in particular, the $2m$ invariant points of a restrictedly elliptic substitution, taken in normal form, consist of m distinct pairs.

It will be noticed that in (5) the $2m$ equations of the substitution are associated in pairs. In its normal form the most general group of the second or elliptic type comprises *all* non-singular elliptic substitutions (5) with the *same* pairing of the coördinates.

1. Derivation of two standard forms for non-loxodromic groups

As remarked in the introduction, we will suppose our projective substitutions to be written throughout in homogeneous form (2) with determinant

$$\Delta \equiv \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} \tag{7}$$

equal to $+1$. Obviously the necessary and sufficient condition that a substitution (2) should be non-loxodromic by the above definition is that the sum of all the principal minors of each order $r \leqq n$ should be real.

When the roots of the characteristic equation (3) are all distinct, the substitution can be transformed through an appropriate substitution T into the canonical form

$$x_i' = \lambda_i x_i \qquad (i = 1, \cdots, n) \qquad (S_1). \tag{8}$$

We will treat only groups of non-loxodromic substitutions which contain at least one such substitution S_1, and we will suppose that the group has been so transformed that S_1 has the form (8). Further, it will be assumed that the subscripts i have been so assigned in (8) that for $i = 1, 2, \cdots, 2m \leqq n$ the λ_i are successively m pairs of conjugate imaginaries, while the remaining λ_i are real.

Consider now any non-loxodromic substitution S, given by (2), which combines with S_1 so as to generate a non-loxodromic group. The sum of principal minors $a_{11} + a_{22} + \cdots + a_{nn}$ is real, likewise the corresponding sum

$$\lambda_1^k a_{11} + \lambda_2^k a_{22} + \cdots + \lambda_n^k a_{nn} \equiv r_k$$

for SS_1^k. Denoting these real sums by r_k and taking $k = 0, 1, \cdots, n-1$, we have a system of n equations linear in the a_{ii}. The determinant of the

system is the product $\Pi(\lambda_i - \lambda_j)$ which cannot vanish in consequence of our hypothesis of the distinctness of the roots of S_1. There is therefore a unique solution for the a_{ii} in terms of the λ_i, r_i. The first $2m$ coefficients λ_i^k in each equation are successively pairs of conjugate imaginaries while the remaining ones are real. If, therefore, we exchange each of the first m odd terms on the left side of each equation with the following even term, we have a system of equations which is the conjugate of the preceding system, a_{ii} and $a_{i+1,\ i+1}$ being interchanged for odd $i < 2m$. Consequently a_{ii} and $a_{i+1,\ i+1}$ are conjugate imaginaries for odd values of $i < 2m$, while the a_{ii} are real for $i > 2m$.

Consider next the combination of (2) with any substitution

$$x_i' = b_{i1}x_1 + b_{i2}x_2 + \cdots + b_{in}x_n \qquad (S') \tag{9}$$

of the same non-loxodromic group. The product $S'S$ is a substitution

$$x_i' = B_{i1}x_1 + B_{i2}x_2 + \cdots + B_{in}x_n \qquad (i = 1, \cdots, n),$$

in which

$$B_{ij} = b_{i1}a_{1j} + b_{i2}a_{2j} + \cdots + b_{in}a_{nj}. \tag{10}$$

Suppose first $m \geqq 1$. Then B_{11}, B_{22} are conjugate imaginaries, and the same is true of the corresponding coefficients

$$\begin{aligned} B_{11}^{(k)} &= \lambda_1^k b_{11}a_{11} + \lambda_2^k b_{12}a_{21} + \cdots + \lambda_n^k b_{1n}a_{n1}, \\ B_{22}^{(k)} &= \lambda_1^k b_{21}a_{12} + \lambda_2^k b_{22}a_{22} + \cdots + \lambda_n^k b_{2n}a_{n2} \end{aligned} \tag{11}$$

in $S'S_1^kS$. Take $k = 0, 1, \cdots, n-1$. We have then in (11) two systems of equations linear in the $b_{1i}a_{i1}$ and $b_{2i}a_{i2}$ respectively with a common determinant $\Pi(\lambda_i - \lambda_j)$ not zero. In the two systems corresponding coefficients λ_i^k are the same, the first m pairs being conjugate imaginaries and the remaining ones real, while the corresponding left members are conjugate imaginaries. We may therefore regard the two systems of equations as conjugate imaginary systems in which the $b_{1i}a_{i1}$ and $b_{2i}a_{i2}$ correspond to one another for $i > 2m$, while for odd $i < 2m$ they correspond respectively to $b_{2,\ i+1}a_{i+1,\ 2}$ and $b_{1,\ i+1}a_{i+1,\ 1}$ respectively. The solution of the two systems therefore gives

$$b_{1i}a_{i1} = \bar{b}_{2,\ i+1}\bar{a}_{i+1,\ 2} \qquad (\text{odd } i < 2m), \tag{12}$$

$$b_{2i}a_{i2} = \bar{b}_{1,\ i+1}\bar{a}_{i+1,\ 1} \qquad (\text{odd } i < 2m), \tag{13}$$

$$b_{1i}a_{i1} = \bar{b}_{2i}\bar{a}_{i2} \qquad (i > 2m). \tag{14}$$

We will now make the *Fundamental Hypothesis* that for each value of $i \leqq n$ there is in the non-loxodromic group at least one substitution (2) for which a_{1i} is not zero and also one for which a_{i1} is not, leaving for later con-

sideration the geometrical significance of this hypothesis. Choose then two substitutions S, S' for which $b_{21}\,a_{12} \neq 0$. Then by (13) b_{12} and a_{21} are not zero. Taking next S' identical with S, we see by (13) that $a_{12}\,a_{21}$ is its own conjugate and hence is real. A very different outcome is obtained according as we suppose that this product is positive or negative. The two corresponding cases we will refer to as *case* 1 *and case* 2 *respectively.*

To normalize the groups in these two cases it will suffice to make a simple change of variable

$$x_i = \rho_i\, y_i, \qquad x_i' = \rho_i\, y_i' \qquad (i = 1, \cdots, n) \tag{15}$$

with a proper determination of the ρ_i, restoring for convenience the letter x in place of y after the transformation. This change does not affect our canonical substitution S_1 nor the values of the a_{ij} in S, while the new value of a_{ij} differs from the old value by the factor ρ_j/ρ_i. Also the values of the principal minors of (7) and the characteristic equation (3) remain unaltered. We will so choose ρ_1, ρ_2 that in the substitution S of the preceding paragraph, for which $a_{12}\,a_{21} \neq 0$, we will have $a_{12} = \bar{a}_{21}$ in case 1 and $a_{12} = -\bar{a}_{21}$ in case 2. It follows then from (13) that we have simultaneously $b_{12} = \bar{b}_{21}$ in case 1, $b_{12} = -\bar{b}_{21}$ in case 2 for every substitution S' of the group.

Next take any odd value of i between 1 and $2m$ and choose two substitutions S, S' for which $b_{1i}\,a_{i1}$ is not zero. Then in consequence of (12) $a_{i+1,\,2}$ and $b_{2,\,i+1}$ are not zero. By properly choosing ρ_{i+1}/ρ_i we may make a_{i1} and $a_{i+1,\,2}$ in S the conjugates of one another. If then S' denotes any substitution whatsoever in the group (including S itself), it follows from (12) that for all substitutions S' we have simultaneously $b_{1i} = \bar{b}_{2,\,i+1}$ for odd $i < 2m$. Taking an S' for which $b_{1i} \neq 0$, we see again from (12) that for every substitution S of the group we have $a_{i1} = \bar{a}_{i+1,\,2}$ for odd $i < 2m$.

Consider next the two coefficients in $S'S_1^kS$

$$B_{12}^{(k)} = \lambda_1^k\, b_{11}\, a_{12} + \lambda_2^k\, b_{12}\, a_{22} + \cdots + \lambda_n^k\, b_{1n}\, a_{n2},$$
$$B_{21}^{(k)} = \lambda_1^k\, b_{21}\, a_{11} + \lambda_2^k\, b_{22}\, a_{21} + \cdots + \lambda_n^k\, b_{2n}\, a_{n1}.$$

These are conjugate imaginaries in case 1 and negative conjugates in case 2. If we set $k = 0, 1, \cdots, n-1$, we have two systems of equations linear in $b_{1i}\,a_{i2}$ and $b_{2i}\,a_{i1}$ respectively. Reasoning in exactly the same manner as above for equations (11), we obtain the following relations between any two substitutions of our group:

$$\begin{aligned} b_{1i}\,a_{i2} &= \bar{b}_{2,\,i+1}\,\bar{a}_{i+1,\,1} && \text{(case 1)} \\ b_{1i}\,a_{i2} &= -\,\bar{b}_{2,\,i+1}\,\bar{a}_{i+1,\,1} && \text{(case 2)} \end{aligned} \qquad (\text{odd } i < 2m) \tag{16}$$

$$\begin{aligned} b_{2i}\,a_{i1} &= \bar{b}_{1,\,i+1}\,\bar{a}_{i+1,\,2} && \text{(case 1)} \\ b_{2i}\,a_{i1} &= -\,\bar{b}_{1,\,i+1}\,\bar{a}_{i+1,\,2} && \text{(case 2)} \end{aligned} \qquad (\text{odd } i < 2m) \tag{17}$$

$$\begin{aligned} b_{1i}\,a_{i2} &= \bar{b}_{2i}\,\bar{a}_{i1} && \text{(case 1)} \\ b_{1i}\,a_{i2} &= -\,\bar{b}_{2i}\,\bar{a}_{i1} && \text{(case 2)} \end{aligned} \qquad (i > 2m) \tag{18}$$

For odd $i < 2m$ our normalization has already made $a_{i1} = \bar{a}_{i+1,\,2}$, $\bar{b}_{1i} = b_{2,\,i+1}$. If, therefore, we select for (17) a substitution S in which the former pair is not zero and for (16) a substitution S' in which the latter pair is not, we get from these equations $b_{2i} = \bar{b}_{1,\,i+1}$, $a_{i2} = \bar{a}_{i+1,\,1}$ in case 1, and $b_{2i} = -\bar{b}_{1,\,i+1}$, $a_{i2} = -\bar{a}_{i+1,\,1}$ in case 2. These results apply to any substitutions S', S in our group. Lastly, if we take a pair of substitutions with $b_{1i}a_{i1} \neq 0$ and divide (18) by (14), we get

$$\frac{a_{i2}}{a_{i1}} = \frac{\bar{a}_{i1}}{\bar{a}_{i2}} \qquad (i > 2m, \text{ case 1}), \tag{19}$$

while in case 2 these two ratios must be the negative of each other. But the latter is impossible, since a quantity cannot equal the negative of the conjugate of its reciprocal. We thus reach the important conclusion that *case 2 is possible only when* $i = 2m$.

In case 1 equation (19) makes $|a_{i1}| = |a_{i2}|$. We have not yet disposed of ρ_i for $i > 2m$, and we will now so take it that the arguments of a_{i1} and a_{i2} in the particular S under consideration shall be the negative of each other. Then we have $\bar{a}_{i1} = a_{i2}$, whence it follows by (18) that in case 1 *we have* $b_{2i} = \bar{b}_{1i}$ $(i > 2m)$ *for every substitution* S' *in the group*. By taking next an S' with $b_{1i} \neq 0$, it follows again from (18) that for every substitution S in the group we have $a_{i2} = \bar{a}_{i1}$ $(i > 2m)$ in case 1.

We have thus completely determined the character of the first two rows and columns of the coefficients (7) of every substitution in our non-loxodromic group. The character of the remaining rows and columns may now be rapidly fixed as follows. We know that for odd $j < 2m$ the coefficients

$$\begin{aligned} B_{1j}^{(k)} &= \lambda_1^k b_{11} a_{1j} + \lambda_2^k b_{12} a_{2j} + \cdots + \lambda_n^k b_{1n} a_{nj}, \\ B_{2\,,\,j+1}^{(k)} &= \lambda_1^k b_{21} a_{1,\,j+1} + \lambda_2^k b_{22} a_{2,\,j+1} + \cdots + \lambda_n^k b_{2n} a_{n,\,j+1} \end{aligned} \tag{20}$$

in $S'S_1^kS$ are conjugate imaginary in both cases 1 and 2. Putting $k = 0, 1, \cdots, n-1$, we have two systems of equations for the n quantities $b_{1i}a_{ij}$ and $b_{2i}a_{i,\,j+1}$ $(i = 1, \cdots, n)$ respectively. Since λ_i^k, λ_{i+1}^k are conjugate imaginaries for odd $i < 2m$ and λ_i is real for $i > 2m$, it follows in the same manner as before that we have the following pairs of conjugate imaginaries:

$$\begin{aligned} &b_{1i}a_{ij} \text{ and } b_{2,\,i+1}a_{i+1,\,j+1} \\ &b_{2i}a_{i,\,j+1} \text{ and } b_{1,\,i+1}a_{i+1,\,j} \end{aligned} \qquad (\text{odd } j < 2m \text{ and odd } i < 2m)$$

$$b_{1i}a_{ij} \text{ and } b_{2i}a_{i,\,j+1} \qquad (\text{odd } j < 2m \text{ and } i > 2m).$$

Now for odd $i < 2m$ we had already $\bar{b}_{1i} = b_{2,\,i+1}$ in both cases 1 and 2, $b_{1,\,i+1} = \bar{b}_{2i}$ in case 1 and $b_{1,\,i+1} = -\bar{b}_{21}$ in case 2; while for $i > 2m$ we had $\bar{b}_{1i} = b_{2i}$ in case 1. Since also by our *Fundamental Hypothesis* we could select S' so that a prescribed one of these quantities should not be zero, it follows that for every substitution S of our group we have the relations:

Case 1.

$$\begin{aligned} \bar{a}_{ij} &= a_{i+1,\,j+1} \\ \bar{a}_{i,\,j+1} &= a_{i+1,\,j} \end{aligned} \qquad (\text{odd } j < 2m;\ \text{odd } i < 2m)$$

$$\bar{a}_{ij} = a_{i,\,j+1} \qquad (\text{odd } j < 2m;\ \text{odd } i > 2m)$$

Case 2.

$$\begin{aligned} \bar{a}_{ij} &= a_{i+1,\,j+1} \\ \bar{a}_{i,\,j+1} &= -\,a_{i+1,\,j+1} \end{aligned} \qquad (\text{odd } j \text{ and } i)$$

The character of the coefficients has just been completely fixed for case 2 and for the first $2m$ columns of (7) in case 1. *In case* 2 *the determinant* (7) *consists of* m^2 *blocks of form* (6), *as stated in the introduction.*

It remains only to fix in case 1 the law of the coefficients subsequent to the $2m$th column of (7). Now it has already been found that $B_{1j}^{(k)}$ and $B_{2j}^{(k)}$ are conjugates for $j > 2m$, and their values will be given by (20) if the subscript $j+1$ in the second equation is replaced throughout by j. Taking then $k = 0, 1, \cdots, n-1$ in the two equations of (20) and arguing in the same manner as before, we obtain for odd $i < 2m$ the conjugate imaginary pairs $b_{1i}\,a_{ij}$ and $b_{2,\,i+1}\,a_{i+1,\,j}$, $b_{2i}\,a_{ij}$ and $b_{1,\,i+1}\,a_{i+1,\,j}$, while for $i > 2m$ we have as conjugates $b_{1i}\,a_{ij}$ and $b_{2i}\,a_{ij}$. For the given odd $i < m$ a substitution S' can be employed for which the conjugate elements b_{1i}, $b_{2,\,i+1}$ or b_{2i}, $b_{1,\,i+1}$ are not zero, and for $i > 2m$ a substitution for which the conjugates b_{1i}, b_{2i} are not. Hence, in every substitution S of our group, a_{ij} and $a_{i+1,\,j}$ are conjugates for odd $i < 2m$, $j > 2m$, while for $i > 2m$, $j > 2m$ every a_{ij} is self-conjugate and consequently real.

The completed result which has thus been reached in case 1 may now be summed up as follows. On the assumption that S_1 possesses distinct characteristic roots λ_i and at least one pair of roots which are conjugate imaginaries, a normal form for a non-loxodromic group satisfying our *Fundamental Hypothesis* has been obtained in which every substitution has the following structure. Its determinant (7) consists of four sections. The minor common to the first $2m$ rows and columns is made up of m^2 blocks of form

$$\begin{vmatrix} a_{ij} & a_{i,\,j+1} \\ \bar{a}_{i,\,j+1} & \bar{a}_{ij} \end{vmatrix} \qquad (\text{odd } i \text{ and } j),$$

and the minor common to the last $n - 2m$ rows and columns contains only real elements. For $i < 2m$, $j > 2m$ the elements occur in conjugate pairs a_{ij}, $a_{i+1,\,j}$ (odd i), and for $i > 2m$, $j < 2m$ in conjugate pairs a_{ij}, $a_{i,\,j+1}$ (odd j). This normal form corresponds to the form (3′) given in the introduction for $m = 1$. The simple change of variable

$$(21) \qquad \begin{aligned} x_{2i-1} &= y_{2i-1} + iy_{2i}, \qquad x_{2i} = y_{2i-1} - iy_{2i} && (i < m), \\ x_i &= y_i && (i > 2m) \end{aligned}$$

throws the group at once into a second and simpler normal form in which the coefficients are all real.

When $m = 0$, the above derivation of the group fails, but the real normal form just indicated is still valid and can be rapidly derived as follows. By hypothesis, all the characteristic roots λ_i of S_1 are now real as well as distinct, and accordingly in every substitution (2) of the group all the a_{ii} are real. Since the left members of the system obtained by putting $k = 0, 1, \cdots, n-1$ in the first equation of (11) are real, all the $b_{1i}\, a_{i1}$ resulting from the solution of the system must be real. Further, by our *Fundamental Hypothesis*, for each value of i there is some S' for which b_{1i} is not zero, and in this particular S' by a change of variable (15) with the proper value of ρ_i/ρ_1 we may make b_{1i} $(i \neq 1)$ real. It follows then that in every substitution S of the group a_{i1} is real. By choosing next an S having $a_{i1} \neq 0$, it follows again from the reality of $b_{1i}\, a_{i1}$ that every substitution S' of the group has a real b_{1i}. In this wise all the elements of the first row and column of (7) may be made simultaneously real for all substitutions of the group. Turning next to the system of equations obtained by putting $k = 0, 1, \cdots, n-1$ in the first equation of (20), we argue in the same manner that every $b_{1i}\, a_{ij}$ is real and hence every a_{ij}. Thus by the simple change of variable (15) we can throw the group into a normal form in which all of its substitutions are real.

We turn now to consider the significance of our *Fundamental Hypothesis* that for each value of i there is some substitution of the group for which $a_{1i} \neq 0$. Make for a moment the contrary hypothesis and suppose that *for some fixed* j we have $a_{1j} = 0$ in every substitution of the group.* Then in (20) we have $B_{ij}^{(s)} = 0$ for this value of j. Taking $k = 0, 1, \cdots, n-1$, we have a homogeneous system of equations giving $b_{1i}\, a_{ij} = 0$ $(i = 1, \cdots, n)$ for *any* two substitutions S, S' of the group. Hence for any value of i we must have either $b_{1i} = 0$ for every substitution S' of the group or $a_{ij} = 0$ for every substitution S of the group. Suppose now that there are exactly r values of j for which a_{1j} is zero in every substitution of the group. Then for each of these r values of j and for any *other* value of i not one of these r values we must have $a_{ij} = 0$ in every substitution of the group. There will be no loss of generality if we assume that in the last r columns of (7) and in no others we have $a_{1j} = 0$, since this can be attained † by merely assigning the subscripts properly to our n variables x_i. Then all the elements of (7) common to these columns and to the first $n - r$ rows will be zero. Consequently the first $n - r$ equations of (2) contain only the variables $x_1, \cdots, x_{n-r}$. This reveals that the space of $r - 1$ (complex) dimensions (or a point, if $r = 1$) obtained

*If some of the a_{i1} are zero for every substitution of the group, a parallel reasoning applies with like result.

†Unless $a_{11} = 0$ for every substitution, when we use the first and the last $r - 1$ columns in place of the last r columns.

by putting these coördinates equal to zero is carried over into itself by all substitutions of the group. Conversely, it is evident that if this is the case, the elements of (7) common to the last r columns and the first $n - r$ rows must be zero for all substitutions of the group. The significance of our *Fundamental Hypothesis* is therefore to forbid that all substitutions of the group should have such an invariant space when S_1 is taken in form (8).

The hyperplanes $x_i = 0$ are the only invariant hyperplanes of S_1 since the multipliers λ_i in (8) are by hypothesis all distinct. The linear spaces common to $n - r$ of these invariant hyperplanes are also invariant under S_1, and it is readily seen that they are the only spaces of $r - 1$ dimensions invariant under S_1 which are common to $n - r$ hyperplanes. For consider an invariant space which is defined by linearly independent hyperplanes

$$u_i \equiv \sum_{j=1}^{n} c_{ij} x_j = 0 \qquad (i = 1, \cdots, n - r).$$

If we take these planes as our coördinate planes along with r other independent hyperplanes, then S_1 in terms of the new u-coördinates will be specified by n equations of form (2) with the u_i in place of the x_i. In the corresponding determinant (7) the elements common to the last r columns and the first $n - r$ rows must be zero since the space common to the hyperplanes $u_i = 0$ $(i = 1, \cdots, n - r)$ is invariant. Accordingly there will be then $n - r$ linearly independent combinations of $u_1, \cdots, u_{n-r}$ corresponding to $n - r$ characteristic roots λ_i, and these, except for constant factors, must be identical with $n - r$ coördinates x_i which figure in (8). Hence the invariant space common to the hyperplanes $u_i = 0$ $(i = 1, \cdots, n - r)$ can also be specified by putting $n - r$ coördinates x_i equal to zero. We thus see that *our Fundamental Hypothesis is identical with the requirement that the substitutions of our group should not admit a common invariant hyperplane nor an invariant space common to two or more hyperplanes.* This requirement is, of course, unaltered when S_1 with the group is transformed out of its normal form. It corresponds to the condition mentioned in the introduction, that a non-loxodromic group of substitutions (1) should not have a common pole.

Up to the present point the development has involved only the reality of the characteristic equations (3) and in no wise the signs of the determinants (7). The conclusions reached therefore apply not only to groups of non-loxodromic substitutions but *likewise to non-singular groups of substitutions whose characteristic equations are real.* For the latter also, under the limitations specified, there are just two classes of groups, and these can be transformed into the two normal forms which have been given. Further, the totality of substitutions of either form, with or without the restriction of being non-loxodromic, form a group. This is evident in the case of substitutions of the

first kind since the totality of real, non-singular substitutions (2) form a group, and the same holds true if the requirement is added that the determinants of the substitutions shall be positive or that they shall be of value $+1$. In the next section it will be shown that the totality of non-singular substitutions of the second kind also form a group, in this case necessarily non-loxodromic.

2. On the nature of the two classes of non-loxodromic groups

A real projective substitution (2) of determinant Δ can be reduced to a form having the determinant unity by dividing each coefficient in (2) by $\Delta^{1/n}$. If Δ is positive, a real nth root may always be chosen so that the substitution remains in real form, but if Δ is negative, this is only possible when n is odd. A real projective substitution in space of $n-1$ (real or complex) dimensions is therefore necessarily non-loxodromic when the number of dimensions is even, but when it is odd, only if the determinant of the substitution is positive.

In accordance with § 1 the *general* non-loxodromic group of the first class, expressed in normal form, consists of all real substitutions (2) of determinant 1. Hence it comprises all non-loxodromic substitutions which leave invariant the real space of the n homogeneous variables which lies within the complex S_{n-1}. When $n-1$ is even, all real and non-singular projective substitutions of the n variables are included. If the group by transformation through any substitution is thrown out of normal form, the real space is converted into an invariant "chain" for the transformed group.

For the special case $n=2$ it is known that the poles of all hyperbolic and parabolic substitutions in a non-loxodromic group of the first class lie upon the invariant real axis or circle, while the poles of any elliptic substitution of the group lie symmetrically with respect to the real axis or circle. We can easily see that there is a corresponding generalization for greater values of n. Consider, for example, any non-loxodromic group of the first class which contains a substitution S whose characteristic roots are all real and *distinct*. Such a substitution will be called *hyperbolic* since it is the natural generalization of the hyperbolic substitution for $n=2$.* Let the group be transformed into its normal real form. Then the hyperbolic substitution becomes a real substitution (2) of unit determinant. To each of our n distinct characteristic roots λ_i belongs a linear form,

$$Y_i = c_1 x_1 + c_2 x_2 + \cdots + c_n x_n,$$

which undergoes the substitution $Y_i' = \lambda_i Y_i$ in consequence of S. The well known equations for determining the coefficients c_i of the form, given in (23) below, show that the ratios of these coefficients are all real. If we equate

* It will presently appear that no hyperbolic substitution is included in a non-loxodromic group of our second class.

these n forms to zero, we obtain a configuration of n real and distinct hyperplanes, and the n vertices of this configuration are the poles of the hyperbolic substitution. The poles of the substitution are accordingly all real, a conclusion which applies simultaneously to all hyperbolic substitutions of the group. Hence *in a non-loxodromic substitution of the first class the poles of its hyperbolic substitutions all lie in the invariant chain of the group.*

Conversely, *if we have a set of hyperbolic substitutions whose poles lie in a common chain, they generate a non-loxodromic group.* For let the group be so transformed that the chain becomes the real coördinate space. The poles of any hyperbolic substitutions in the group are now real and by a real transformation of the group may be made to be the n-coördinate vertices. The substitution then has the real form $x_i' = \lambda_i x_i$ ($i = 1, \cdots, n$) and hence must have had also a real form before the last transformation. Consequently all hyperbolic substitutions of the group have been made simultaneously real. Now by hypothesis our hyperbolic substitutions are of unit-determinant; they generate therefore a group consisting only of real substitutions of unit-determinant,—that is, a group of non-loxodromic substitutions.

Take now the other extreme case in which the characteristic roots of some particular substitution of a non-loxodromic group of the first class are distinct pairs of conjugate imaginaries. Such a substitution will be called *restrictedly elliptic* and is possible only when n is an even number $2m$. By transformation of the group the particular substitution can be thrown into the form (8) and be used as our substitution S_1 of the preceding section. Without change of S_1 the group was there brought into a first normal form and then was converted into the final real form by the transformation (21). Since the characteristic roots of our substitution are by hypothesis all distinct, the invariant points of (8) were the x-coördinate vertices. The transformation (21) converts the invariant point for which x_{2i-1} is the only non-zero coördinate into a point for which the y-coördinates are all zero except $y_{2i-1} = x_{2i-1}$, $y_{2i} = -ix_{2i-1}$, while the invariant point having x_{2i} as its only non-zero coördinate is converted into one for which all the y-coördinates are zero except $y_{2i-1} = x_{2i}$, $y_{2i} = ix_{2i}$. These two points are imaginary points which lie symmetrically with respect to the real space of the y-coördinates; i.e., the space for which the ratios of the $2m$ homogeneous coördinates are real.* Thus the $2m$ invariant

* If we think of a complex space S_k as a real space with $2k$ real orthogonal axes, two points lie symmetrically with respect to the linear space determined by r-coördinate axes if the corresponding r-coördinates of the two points agree, while the remaining coördinates of one point differ only in sign from those of the other. Accordingly, in the case before us, if we interpret the n complex coördinates y_i as nonhomogeneous coördinates for a real space of $2n$ dimensions, the $n = 2m$ invariant points of our substitution are obviously m pairs of points symmetrically situated with respect to the linear space of n dimensions which consists of the points for which the y_i are *all* real. When, however, as above, the coördinates are thought of as projective, some one of the y_i equated to zero—let us say $y_n = 0$—specifies the region at

points of the substitution S_1 in its final real form are m pairs of points symmetrically situated with respect to this real space of $2m - 1$ dimensions embedded in the complex space of the coördinates having double this number of dimensions. This real space is the invariant real space of the normalized real group in which our substitution is contained. The other "restrictedly elliptic" substitutions of the group, whose characteristic roots are *distinct* pairs of conjugate imaginaries, have become real simultaneously with S_1. Now any two *real* forms into which a given substitution can be transformed are convertible the one into the other by a real linear change of coördinate variables. Such a change in our y-coördinates converts a pair of points, symmetrically situated with respect to the real space of the coördinate-ratios and having corresponding coördinates the conjugates of one another, into another like pair of coördinates. Since, furthermore, any restrictedly-elliptic substitution of the group could be made our S_1, it follows that the invariant points of all restrictedly elliptic substitutions of a normalized non-loxodromic group of the first species consist of pairs of points which lie symmetrically with respect to the real invariant space of the group. If the group is transformed through any substitution T, the real invariant space is converted into a chain, and the images of the invariant point-pairs of these substitutions are converted into pairs of points which may be said to lie symmetrically with respect to this chain.

Turn now to the groups of the second class which are more novel and therefore more interesting. We will first prove that any substitution whose determinant consists of m^2 blocks of form (6) has a real characteristic equation. To establish this it is necessary to show that in its determinant (7) the sum of all the principal minors of any given order r is real. It suffices to prove that the principal minors which are not real occur in conjugate pairs. With this in view let us

(I) exchange each odd row and column of (7) with the following even row and column, and

(II) change first the signs of all the elements in the even rows and then the signs of the elements in the even columns of Δ.

By these two changes every element of (7) is replaced by its conjugate. Since the determinant is unaltered in value by the two changes, it must be real, being identical with its conjugate. Any principal minor of (7) formed

infinity and the nonhomogeneous coördinates will be $n - 1$ ratios $y_1/y_n, \cdots, y_{n-1}/y_n$. The difficulty arising from the presence of pairs of points for which y_n equals zero may be removed by regarding such a pair as the limit of another pair with the same coördinates except y_n, which is replaced by a real number h approaching zero. If, then, for $h \neq 0$ the two points are symmetrically situated with respect to the real space for which the $(n - 1)$-coördinate ratios are real, the limiting points $(h = 0)$ will also be said to lie symmetrically with respect to this space.

by deleting with each odd row and column the following even row and column is real for the same reason. Any other principal minor in consequence of I is exchanged with another principal minor, but neither is changed in sign by II. Hence these two minors must be the conjugate of one another, and the substitution has accordingly a real characteristic equation, as stated.

It will be seen later that a non-zero determinant Δ having the structure (6) must be positive. If therefore the substitution is non-singular, its determinant can be made equal to $+1$ by dividing all its coefficients by a real $\sqrt[n]{\Delta}$. *All non-singular substitutions of form* (5) *are therefore non-loxodromic.*

Consider next the product of any two (non-singular or singular) substitutions S, S' of the second class with coefficients a_{ij} and b_{ij} respectively, and with the same pairing of equations in each, as indicated in (5). Since the coefficients of both obey the law (6), we find at once from (10) that $B_{i+1,\,j+1} = \overline{B}_{ij}$, $B_{i+1,\,j} = -\overline{B}_{i,\,j+1}$ for odd i and j. Consequently the product of the two substitutions has the same form.

Furthermore, the inverse of a non-singular substitution will also be of like form. In fact, since its determinant (7) is real, it is only necessary to see that the determinant of the cofactors A_{ij} of the elements of (7) consists of blocks of the form (6). Now we have just seen that the changes I and II together replace every element of (7) by its conjugate and therefore every A_{ij} by its conjugate. On the other hand, these changes obviously result in an interchange of A_{ij} and $A_{i+1,\,j+1}$ for odd i and j, also of $A_{i+1,\,j}$ and $A_{i,\,j+1}$, with a reversal of sign in the latter case. Hence we have $A_{i+1,\,j+1} = \overline{A}_{ij}$, $A_{i+1,\,j} = -\overline{A}_{i,\,j+1}$. Thus the inverse substitution has the same form, and we conclude that *any number of non-singular substitutions* (5) *of our second kind generate a non-loxodromic group.*

We now proceed to examine the nature of the substitutions of our second class. Let λ_i be a characteristic root of any substitution (2) of the second class. To this there corresponds a linear form

$$Y_1 = c_1 x_1 + c_2 x_2 + \cdots + c_n x_n \qquad (n = 2m), \tag{22}$$

which in consequence of the substitution undergoes the transformation $Y_1' = \lambda_i Y_1$, and its coefficients satisfy the equations

$$\begin{aligned} &c_1(a_{11} - \lambda_i) + c_2 a_{21} + \cdots + c_n a_{n1} = 0, \\ &c_1 a_{12} + c_2(a_{22} - \lambda_i) + \cdots + c_n a_{n2} = 0, \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ &c_1 a_{1n} + c_2 a_{2n} + \cdots + c_n(a_{nn} - \lambda_i) = 0, \end{aligned} \tag{23}$$

the consistency of which is guaranteed by our characteristic equation. Suppose now every quantity which appears in these equations to be replaced by its conjugate. We get a system of equations for $\bar{c}_1, \cdots, \bar{c}_n$. Exchange next

each odd equation in (23) with the following even equation, then each odd column with the following even column, and finally change the signs of the n terms of the resulting even equations. Then by virtue of the fact that a_{ij}, $a_{i+1,\ j+1}$, also $a_{i+1,\ j}$ and $-a_{i,\ j+1}$ are conjugates for odd i and j, we get precisely the equations which say that

$$(24) \qquad -Z_1 = (\bar{c}_2 x_1 - \bar{c}_1 x_2) + (\bar{c}_4 x_3 - \bar{c}_3 x_4) + \cdots + (\bar{c}_n x_{n-1} - \bar{c}_{n-1} x_n)$$

undergoes the transformation $Z_1' = \overline{\lambda}_i Z_1$ in consequence of our substitution (2). Thus the pair of linear forms which correspond to a simple pair of conjugate imaginary roots of the characteristic equation have the form (22) and (24). Moreover, the two forms cannot be linearly dependent even when λ_i is self-conjugate and real, for then we would have

$$\frac{\bar{c}_{2i}}{c_{2i-1}} = -\frac{\bar{c}_{2i-1}}{c_{2i}}, \qquad \text{or} \qquad c_{2i}\,\bar{c}_{2i} + c_{2i-1}\,\bar{c}_{2i-1} = 0 \qquad (i = 1, \cdots, m),$$

which would make every $c_i = 0$ in (22). It follows that every real root of the characteristic equation must be at least a double root.

In case the elementary divisors of the characteristic equation are not all simple, there correspond s linear forms to each elementary divisor $(\lambda - \lambda_i)^s$ of the x_i which in consequence of the substitution undergo the transformation

$$(25) \qquad \begin{aligned} Y_1' &= \lambda_i Y_1, \\ Y_2' &= Y_1 + \lambda_i Y_2, \\ Y_3' &= Y_2 + \lambda_i Y_3, \\ &\cdots\cdots\cdots \\ Y_s' &= Y_{s-1} + \lambda_i Y_s. \end{aligned}$$

Let us suppose that Y_1 is given by the right member of (22), and correspondingly Z_1 by (24). Put

$$(26) \qquad Y_2 = d_1 x_1 + d_2 x_2 + \cdots + d_n x_n \qquad (n = 2m).$$

By use of (1) and (26) in the second equation of (25) we get

$$(27) \qquad \begin{aligned} &d_1(a_{11} - \lambda_i) + d_2 a_{21} + \cdots + d_n a_{n1} = c_1, \\ &d_1 a_{12} + d_2(a_{22} - \lambda_i) + \cdots + d_n a_{n2} = c_2, \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ &d_1 a_{1n} + d_2 a_{2n} + \cdots + d_n(a_{nn} - \lambda_i) = c_n. \end{aligned}$$

Replace every quantity which appears in these equations by its conjugate, exchange each odd equation with the following even equation, then exchange each odd column with the following even column, and finally change the signs of the terms on both sides of the even equations. We get a set of equations which state that the linear form

$$(28) \qquad -Z_2' = (\bar{d}_2 x_1 - \bar{d}_1 x_2) + \cdots + (\bar{d}_n x_{n-1} - \bar{d}_{n-1} x_n)$$

undergoes the transformation

$$Z_2' = Z_1 + \bar{\lambda}_i Z_2$$

in consequence of our substitution. Proceeding next in the same way with the third equation of (25) and so on in turn with each of its succeeding equations, we find that to the s linear forms $Y_1, \cdots, Y_s$ there correspond also s other linear forms $Z_1, \cdots, Z_s$, which are related to them in exactly the same way as (24) to (22) and which undergo the transformation

$$\begin{aligned} Z_1' &= \bar{\lambda}_i Z_1, \\ Z_2' &= Z_1 + \bar{\lambda}_i Z_2, \\ &\cdots\cdots\cdots \\ Z_s' &= Z_{s-1} + \bar{\lambda}_i Z_s, \end{aligned} \tag{29}$$

in consequence of the substitution. The coexistence of (29) and (25) indicates that for each elementary divisor $(\lambda - \lambda_i)^s$ there is another divisor $(\lambda - \bar{\lambda}_i)^s$, and hence gives the following result:

THEOREM. *The elementary divisors of the characteristic equation of any non-singular substitution* (5) *occur only in conjugate pairs* $(\lambda - \lambda_i)^s$, $(\lambda - \bar{\lambda}_i)^s$ *of equal degree, this conclusion holding also for the real characteristic roots* λ_i *to which belong pairs of equal elementary divisors.*

Since the determinant of a substitution is the product of the characteristic roots, taken each with its proper multiplicity, the theorem just stated gives at once the following important corollary:

COROLLARY. *If a non-singular substitution has the form* (5), *its determinant is positive.*

Determinants with *real* elements satisfying the condition (6) have been previously encountered in some work of Voigt and have been called *Determinants of Voigt* by Kowalewski. The proofs* given to show that Voigt's determinant is positive or zero fail altogether to cover the more general determinants here considered in which the elements are complex. As the proof furnished above was indirect, I shall also give a simple and direct proof.

The non-negative character of the determinant is obvious when $m = 1$ since it is then $a_{11}\bar{a}_{11} + a_{12}\bar{a}_{12}$. The extension to any other value of m is obtained in the following manner by mathematical induction. Suppose it to have been established for any value of m that all determinants of order $2m - 2$ having the structure (6) are non-negative. Take then any de-

* Kowalewski, *Einführung in die Determinantentheorie*, § 69, p. 156. Baltzer's and Drude's proofs are both found in the Göttinger Nachrichten, 1887.

Drude and Kowalewski prove, when the elements of (7) are real, that Δ is the sum of the squares of two expressions. When the elements are complex, the natural extension would be to replace each square by the product of the expression and its conjugate. But the resulting sum no longer represents Δ.

terminant (7) of order $2m$, which obeys the same law, and multiply it by *any* minor Δ_{2m-2} of order $2m-2$ obtained by deleting from Δ any odd row and any odd column and the following even row and column. This minor has the structure (6), and there will be no loss of generality in supposing that it consist of the first $2m-2$ rows and columns of Δ since the deleted rows and columns may be moved to the bottom and to the right-hand border of Δ without violating the law of its formation. By a well known identity* we then have

$$\Delta_{2m-2}\Delta = A_{2m-1,\,2m-1}A_{2m,\,2m} - A_{2m-1,\,2m}A_{2m,\,2m-1}. \tag{30}$$

Now we have already observed that $A_{2m-1,\,2m-1}$ and $A_{2m,\,2m}$ are conjugates, while $A_{2m-1,\,2m}$ and $A_{2m,\,2m-1}$ are negative conjugates. Hence the right member of (30) is non-negative, being positive unless all four A_{ij} on the right of (30) are zero. If, therefore, a single one of the minors Δ_{2m-2} formed in the manner indicated is positive, it follows that Δ is positive or zero. On the other hand, if every minor Δ_{2m-2} is zero, every A_{ij} by (30) must be zero and hence also Δ itself. Consequently, *any determinant consisting of blocks of form* (6) *must be non-negative*. Hence the determinant of any non-singular substitution of form (5) is positive.

Any non-singular substitution in n variables which consists only of pairs of components of form (25) and (29) is a substitution of our second class since it can be thrown into form (5) by alternating the equations of (29) with those of (25). Thus, for example, the alternation of their first two equations gives a matrix of coefficients

$$\begin{Vmatrix} \lambda & 0 & 0 & 0 \\ 0 & \bar{\lambda} & 0 & 0 \\ 1 & 0 & \lambda & 0 \\ 0 & 1 & 0 & \bar{\lambda} \end{Vmatrix}$$

having the structure (6).

In establishing the existence of the linear forms Y_i, Z_i for any non-singular substitution of the second class, we implicitly reduced it to canonical form. If we regard our m pairs of equations of form (22) and (24) as a change of variable, the substitution in terms of the new variables consists of pairs of components (25) and (29). Now the transforming equations (22) and (24) are of exactly the same form as (5). Hence any substitution of the second class can be thrown into canonical form by transformation through an appropriately chosen transformation of the same class. Since all the transformations (5) form a group, this can be also done without destroying the form (5) of a non-loxodromic group in which it is contained.

An important property of substitutions (5) *of our second class is to convert a pair of points*

$$P = (x_1, x_2; \cdots; x_{2m-1}, x_{2m}), \qquad Q = (-\bar{x}_2, \bar{x}_1; \cdots; -\bar{x}_{2m}, \bar{x}_{2m-1})$$

* Muir's *Theory of Determinants*, § 90.

into another similar pair. By direct substitution in (5) it will be found at once that when the substitution converts P into $P' = (x'_{2i-1}, x'_{2i}), i = 1, \cdots, m$, it also converts Q into $Q' = (-\bar{x}'_{2i}, \bar{x}'_{2i-1})$. A pair of such points we will speak of as *conjugate* with respect to the substitution or group.

We will now prove, conversely, that *if any substitution* (2) *in our* $2m$ *variables has the property of carrying every pair of conjugate points into a pair of conjugate points, it must be one of our second class having the form* (5). Consider the first two equations of (2) which, for convenience, we will write

$$x'_1 = a_1 x_1 + a_2 x_2 + \cdots + a_n x_n,$$
$$x'_2 = b_1 x_1 + b_2 x_2 + \cdots + b_n x_n \qquad (n = 2m).$$

The coefficients of either equation will, in general, be determined by the requirement that m pairs of conjugate points P_i, Q_i shall be transformed into conjugate pairs P'_i, Q'_i. We get thus the following sets of equations for the determination of our coefficients a_i, b_i:

$$x'_{i1} = a_1 x_{i1} + a_2 x_{i2} + \cdots + a_{2m-1} x_{i,\,2m-1} + a_{2m} x_{i,\,2m},$$
$$-\bar{x}'_{i2} = -a_1 \bar{x}_{i2} + a_2 \bar{x}_{i1} + \cdots - a_{2m-1} \bar{x}_{i,\,2m} + a_{2m} \bar{x}_{i,\,2m-1}$$
$$(i = 1, \cdots, m),$$

$$x'_{i2} = b_1 x_{i1} + b_2 x_{i2} + \cdots + b_{2m-1} x_{i,\,2m-1} + b_{2m} x_{i,\,2m},$$
$$\bar{x}'_{i1} = -b_1 \bar{x}_{i2} + b_2 \bar{x}_{i1} + \cdots - b_{2m-1} \bar{x}_{i,\,2m} + b_{2m} \bar{x}_{i,\,2m-1}$$
$$(i = 1, \cdots, m).$$

The determinants of these two sets of equations are identical and must be positive or zero since they are composed of blocks of form (6). To exclude the value zero, it will be assumed that m pairs of conjugate points P, Q are taken which do not lie in a common hyperplane of our complex S_{n-1}. The numerators of the values obtained for a_1 and b_2 are determinants differing only in their first two columns which consist of m blocks

$$\begin{Vmatrix} x'_{i1} & x_{i2} \\ -\bar{x}'_{i2} & \bar{x}_{i1} \end{Vmatrix}, \qquad \begin{Vmatrix} x_{i1} & x'_{i2} \\ -\bar{x}_{i2} & \bar{x}'_{i1} \end{Vmatrix},$$

respectively, while the corresponding blocks in the numerators of a_2 and b_1 are

$$\begin{Vmatrix} x_{i1} & x'_{i1} \\ -\bar{x}_{i2} & -\bar{x}'_{i2} \end{Vmatrix}, \qquad \begin{Vmatrix} x'_{i2} & x_{i2} \\ \bar{x}'_{i1} & \bar{x}_{i1} \end{Vmatrix}.$$

The remaining pairs of columns are identical for all four numerators and are of structure (6). If now in the numerator of b_2 we exchange each odd row and column with the following even row and column and change first the signs of the even columns and then of the even rows, each element of the numerator determinant is converted into the conjugate of the corresponding

element in the numerator of a_1. Consequently $b_2 = \bar{a}_1$. A like manipulation of the numerator of b_1 gives $b_1 = -\bar{a}_2$. In the same way we obtain $b_{2i} = \bar{a}_{2i+1}$, $b_{2i+1} = -\bar{a}_{2i}$. Hence the first two equations of the substitution must have the structure (6). The reasoning applies to all its succeeding pairs of equations. *Thus the property of converting pairs of conjugate points into conjugate pairs completely characterizes the second class of substitutions.*

One immediate corollary of this result is that the invariant points of a substitution of the second class consist of pairs of conjugate points, for clearly if one point of a conjugate pair is invariant, the other must be. The geometrical interpretation of this is well known for $m = 1$. If, namely, complex coördinates are given to the points of a sphere in the proper way, the formula (4) represents a rotation of the sphere about the diameter passing through the two invariant points of the sphere, and the coördinates* of any two diametrically opposite points are z, $-1/\bar{z}$, or, in homogeneous notation, (z_1, z_2) and $(-\bar{z}_2, \bar{z}_1)$.

It has thus appeared that *our second class of groups generalizes the characteristic form and property of the familiar elliptic rotation groups for* $m = 1$. For $m > 1$ it accordingly seems fitting to apply the term *elliptic* to any substitutions of our second kind, that is, to the substitutions whose elementary divisors occur only in conjugate imaginary or duplicate real pairs. The general elliptic group will then consist of all non-singular substitutions of the second class *with the same pairing of the coördinates*, as indicated in equation (5). A subclass will comprise the restrictedly elliptic substitutions whose characteristic roots are *n distinct* pairs of conjugate imaginaries.

* Cf. Klein's *Vorlesungen über das Ikosaeder*, pp. 32–34.

UNIVERSITY OF WISCONSIN,
MADISON, WIS.

BIRATIONAL TRANSFORMATIONS SIMPLIFYING SINGULARITIES OF ALGEBRAIC CURVES*

BY

GILBERT AMES BLISS

In a preceding paper† I have commented on the proofs of the theorem that every irreducible algebraic curve can be transformed by a birational transformation into one having no singularities except double points with distinct tangents. There are many proofs, two of which seem to be especially interesting. The first is by Walker,‡ who has worked out in detail an alteration, suggested by Klein in 1894,§ of a proof originally devised by Bertini for the projective plane. In the second, by Hensel and Landsberg,¶ reasoning proposed by Kronecker in 1881|| is extended to apply to the proof of the theorem in the function-theoretic plane. These two proofs appeal to me as being the best ones which I know for the two cases, but both of them are lengthy and complicated when all details are taken into consideration. In the present paper I have extended once more the method of Kronecker so that it can be applied in both the function-theoretic and the projective planes, and have attained what I hope will be regarded as simpler proofs of the two corresponding theorems.

It should perhaps be said that Kronecker's original reduction of the singularities of an algebraic curve to ordinary double points was effective only for the finite part of the plane. The alteration of his method to include the points at infinity in the function-theoretic plane is a real extension, as one may see by a study of the proof of Hensel and Landsberg. Still more is required of the transformation if higher singularities are to be excluded from points on the curve in the infinite region as well as in the finite part of the projective plane.

In Section 1 below a set of properties of an algebraic function $y(x)$ defined by an irreducible algebraic equation $f(x, y) = 0$ is described, such that when

* Presented to the Society, April 28, 1923.

† *The reduction of singularities of plane curves by birational transformation*, Bulletin of the American Mathematical Society, vol. 29 (1923), pp. 161–183.

‡ *On the resolution of higher singularities of algebraic curves into ordinary nodes*, Dissertation, Chicago, 1906.

§ See a footnote to the paper by Bertini, *Trasformazione di una curva algebrica in un'altra con soli punti doppi*, Mathematische Annalen, vol. 44 (1894), p. 158.

¶ *Theorie der algebraischen Funktionen*, p. 402, § 2; see also Hensel, *Encyclopädie der mathematischen Wissenschaften*, II C 5, § 25.

|| Journal für die reine und angewandte Mathematik, vol. 91 (1881), p. 301.

the function $y(x)$ has these properties the corresponding algebraic curve $f = 0$ has no singularities in the projective plane except double points with distinct tangents. For the simplification of the singularities of an arbitrary algebraic curve $f(x, y) = 0$ the method followed in the succeeding sections consists in the construction of a pair of rational functions $\xi(x, y)$, $\eta(x, y)$, satisfying an irreducible equation $\phi(\xi, \eta) = 0$ birationally related to $f(x, y) = 0$, and having the properties described in Section 1. In Sections 2, 3, 4 the construction of ξ and η is described and their properties determined. The lemma in Section 2 is particularly effective in simplifying the proofs. In Section 5 the method of the preceding sections is applied in the function-theoretic plane, the general plan being that of Hensel and Landsberg, but the details quite different. In Section 6 the distinction between the theorems for the projective and the function-theoretic planes is explained in terms of homogeneous coördinates.

1. **A preliminary lemma.** Consider an equation $f(x, y) = 0$ in which f is an irreducible polynomial in x and y of degree n in y. Near a finite value $x = a$ the values of y which satisfy this equation are defined by a number of series in increasing integral powers of t of the form

$$y = bt^{\mu} + b't^{\mu'} + \cdots, \tag{1}$$

where $t = (x - a)^{1/r}$, r is a positive integer, and the numbers $r, \mu, \mu', \cdots$ are relatively prime. Each such series defines r values of y when the r values of $(x - a)^{1/r}$ are substituted, and the sum of the integers r for the various series in powers of $(x - a)$ is equal to n. There are only a finite number of values $x = a$, the branch values, for which some of the corresponding exponents r are greater than unity. The polynomial $\Delta(x)$ is defined to be the product of the factors $(x - a)^{r-1}$ corresponding to the series (1) which have $r > 1$,

$$\Delta(x) = \Pi(x - a)^{r-1}.$$

Near $x = \infty$ the values of y are given by expansions (1) in which $t = (1/x)^{1/r}$.

Lemma 1. *An irreducible algebraic curve $f(x, y) = 0$ of degree n in y has no singularities in the projective xy-plane except double points with distinct tangents, provided that the algebraic function $y(x)$ defined by the equation has the following properties*:

1. *at $x = \infty$ the n values of $y(x)$ are given by n expansions*

$$y = x\left(b_k + b_k'\frac{1}{x} + \cdots\right) \quad (k = 1, \cdots, n; b_k \neq b_l \text{ if } k \neq l) \tag{2}$$

and $y(x)$ has no other poles;

2. *the discriminant of $y(x)$ has the form*

$$D(x) = V(x)^2 \Delta(x)$$

where $V(x)$ is a polynomial with roots distinct from each other and from those of $\Delta(x)$.

To prove this lemma let $f(x, y)$ be first expanded in powers of x, y in the form

$$f(x, y) = (x, y)_0 + (x, y)_1 + \cdots + (x, y)_\nu,$$

where $(x, y)_k$ is homogeneous and of degree k in x and y. From the assumption that the series (2) satisfy $f = 0$ it follows readily that $(1, b_k)_\nu = 0$ $(k = 1, \cdots, n)$, and therefore $\nu \geqq n$. It is clear that $(x, y)_\nu$ is exactly of degree n in y since f itself has by hypothesis this property, and furthermore that $\nu = n$, since otherwise $f(x, y)$ would contain y^n multiplied by a polynomial in x, and the function $y(x)$ would necessarily have poles at finite values of x, contrary to the property 1. The intersections of $f = 0$ with the line at infinity are defined by the equation $(x, y)_\nu = 0$ and are evidently distinct. None of them is a multiple point of the curve.

At a finite value $x = a$ distinct from the roots of $D(x)$ every root b of $f(a, y)$ is simple and $f = 0$ has no multiple point at (a, b).

If the roots of $f(x, y)$ corresponding to a value x are represented by y_i $(i = 1, \cdots, n)$, the discriminant has the value

$$D(x) = V(x)^2 \Delta(x) = \prod_{i<j} (y_i - y_j)^2 \quad (i, j = 1, \cdots, n). \tag{3}$$

For a root $x = a$ of $\Delta(x)$ the n values of y are defined by expansions (1) in positive integral powers of $(x - a)^{1/r}$ which have the form

$$y = b + b'(x - a)^{1/r} + \cdots, \tag{4}$$

since $y(x)$ has no pole at a finite value of x. Each of these series furnishes $r(r-1)/2$ differences $y_i - y_j$ of order at least $1/r$ in $(x - a)$. Since each factor $y_i - y_j$ occurs twice in $D(x)$ it follows that for each such series $D(x)$ contains at least the factor $(x - a)^{r-1}$, whatever the values of b and b' may be. Since, by property 2, $D(x)$ can contain no more powers of $(x - a)$ than are contained in $\Delta(x)$, it follows readily that for each series (4) corresponding to $x = a$ the coefficient b' must be different from zero, and that the coefficients b for the different series corresponding to $x = a$ must be distinct. The value b is therefore exactly an r-tuple root of $f(a, y)$, and $(y - b)^r$ is the lowest power of $(y - b)$ alone actually occurring in the expansion

$$f(x, y) = (x - a, y - b)_1 + \cdots + (x - a, y - b)_n \tag{5}$$

of $f(x, y)$ at (a, b). It may also be verified that the expansion (4) can satisfy $f = 0$ only if the term in $(x - a)$ is present in $(x - a, y - b)_1$. Hence none of the points (a, b) corresponding to a root $x = a$ of $\Delta(x)$ is a singular point of the curve $f = 0$.

At a zero $x = a$ of $V(x)$ the expansions (4) all have $r = 1$, since the roots of $V(x)$ and $\Delta(x)$ are distinct by the property 2 of the lemma. Since, by the property 2, $D(x)$ has exactly the factor $(x - a)^2$, it follows that one only of the differences $y_i - y_j$ has a simple zero at $x = a$, and that for it the expansions of y_i and y_j have the form

$$y_i = b + c_1(x - a) + \cdots, \qquad y_j = b + c_2(x - a) + \cdots, \tag{6}$$

with $c_1 \neq c_2$. The value $y = b$ is a double root of $f(a, y)$, and $(y - b)^2$ is the lowest power of $(y - b)$ alone actually occurring in the expansion (5). The series (6) can make (5) vanish identically only if the remaining linear term in $(x - a)$ is absent and if $(1, c_1)_2 = (1, c_2)_2 = 0$. Hence (a, b) is a double point of the curve $f = 0$ with distinct tangents. The other roots b' of $f(a, y)$ are in this case simple, and the points (a, b') are simple points of the curve.

2. **Rational functions $\sigma(x, y)$ with simple poles.** When the series (1) and the value of x in terms of t are substituted in a rational function $\sigma(x, y)$ of the variables x, y, an expansion for σ is found of the form

$$\sigma(x, y) = ct^{\mu} + dt^{\mu+1} + \cdots. \tag{7}$$

The constant μ is a positive or negative integer, not necessarily the same as the integer μ in (1), and is called the order of $\sigma(x, y)$ at the expansion (1). On the Riemann surface T of $f(x, y) = 0$ the places P are in one-to-one correspondence with the expansions (1). Hence μ is also called the order of $\sigma(x, y)$ at P.

A symbol $\mathbf{D} = P_1^{\mu_1} \cdots P_s^{\mu_s}$, where the P's represent places on the Riemann surface and the μ's are positive or negative integers, is called a divisor. The sum $\mu_1 + \cdots + \mu_s = -q$ is the order of the divisor. A rational function $\sigma(x, y)$ is a multiple of the divisor $\mathbf{D}$ if its orders μ are $\geqq \mu_i$ at the places P_i $(i = 1, \cdots, s)$ and $\geqq 0$ elsewhere. One of the most important problems of the theory of algebraic functions* is the determination of the functions $\sigma(x, y)$ which are multiples of a divisor $\mathbf{D}$. It is a theorem that if $q > 2p - 2$, where p is the genus of the curve $f(x, y) = 0$, then the number of linearly independent multiples of $\mathbf{D}$ is exactly $\nu = q - p + 1$.†

* Hensel, *Encyclopädie der mathematischen Wissenschaften*, II C 5, p. 557.

† Consider two divisors Q, Q' with orders q, q', and whose product is in the differential class (W). Let $\{Q\}$, $\{Q'\}$ represent the numbers of linearly independent multiples of $1/Q$ and $1/Q'$, respectively. Then, by the Riemann-Roch theorem,

$$q + q' = 2p - 2, \qquad \{Q'\} = \{Q\} - q + p - 1.$$

See Hensel und Landsberg, *Theorie der algebraischen Funktionen*, pp. 301, 304, formulas (1*a*) and (II); and Hensel, loc. cit., pp. 557–8. If $q > 2p - 2$ then $q' = 2p - 2 - q < 0$, and the number $\{Q'\}$ of multiples of $1/Q'$ is zero. Hence in this case $\{Q\} = q - p + 1$. In the text above the $1/Q$ of this footnote is replaced by $\mathbf{D}$.

Consider now a divisor $\mathsf{D} = P_1^{-1} \cdots P_q^{-1}$ for which $q > 2p + 2$. The $\nu = q - p + 1$ linearly independent multiples $\sigma_1, \cdots, \sigma_\nu$ of this divisor are everywhere finite except possibly for simple poles at the places in D. The coefficient c of the lowest power permitted in the expansion (7) for a multiple of D may be called the leading coefficient of σ at the corresponding place P, and for the functions $\sigma_1, \cdots, \sigma_\nu$ these coefficients will be denoted by $c_1, \cdots, c_\nu$.

LEMMA 2. *For the multiples* $\sigma_1, \cdots, \sigma_\nu$ *of a divisor* $\mathsf{D} = P_1^{-1} \cdots P_q^{-1}$ $(q > 2p + 2)$ *the matrix*

$$\begin{Vmatrix} c_1 & \cdots & c_\nu \\ c_1' & \cdots & c_\nu' \\ c_1'' & \cdots & c_\nu'' \\ c_1''' & \cdots & c_\nu''' \end{Vmatrix} \tag{8}$$

of leading coefficients at four arbitrarily selected places P, P', P'', P''' *on the Riemann surface of* $f(x, y) = 0$ *has always rank* 4.

Suppose that this were not true for a set of places P, P', P'', P'''. Then there would be at least $\nu - 3$ linearly independent functions of the form $v_1\sigma_1 + \cdots + v_\nu\sigma_\nu$ with $v_1, \cdots, v_\nu$ constants and with leading coefficients zero at P, P', P'', P'''. These functions would be multiples of the divisor $\mathsf{D}_1 = PP'P''P'''\mathsf{D}$ of order $4 - q$. But this is impossible, since $q - 4 > 2p - 2$ and the number of linearly independent multiples of D_1 is therefore

$$q - 4 - p + 1 = \nu - 4.$$

COROLLARY 1. *The matrix corresponding to* (8) *for* $k < 4$ *places has always the rank* k.

If the place P' is replaced by P, and the second row in the matrix (8) by the second coefficients $d_1, \cdots, d_\nu$ of the expansion (7) for $\sigma_1, \cdots, \sigma_\nu$, then the lemma and its proof would still be valid. The following corollary is therefore also true:

COROLLARY 2. *At every place* P *the second coefficients* $d_1, \cdots, d_\nu$ *in the expansions of* $\sigma_1, \cdots, \sigma_\nu$ *are not all zero.*

3. **Properties of pairs of rational functions.** The properties described in this section for pairs of rational functions $\xi(x, y)$, $\eta(x, y)$ on the Riemann surface T of an irreducible algebraic equation $f(x, y) = 0$ are well known.* They are reproduced here, however, in a form which is especially adapted to the applications which are to be made in Section 4.

In the first place if $\xi(x, y)$ has q poles on the surface T, then for every constant α the function $\xi(x, y) - \alpha$ will have q zeros, and it is clear that at the q places of T where $\xi(x, y)$ has a fixed value the function $\eta(x, y)$ takes q values $\eta_1, \cdots, \eta_q$ which may or may not be distinct. To obtain analytic ex-

* Hensel und Landsberg, loc. cit., p. 247. Appel and Goursat, *Théorie des Fonctions algébriques*, p. 256.

pressions for $\eta_1, \cdots, \eta_q$ for values of ξ near $\xi = \alpha$, consider the expansions analogous to (7) for the functions $\xi(x, y) - \alpha$ and $\eta(x, y)$ near a zero of the former. These have the form

$$(9)\quad \xi - \alpha = ct^\rho + dt^{\rho+1} + \cdots, \quad \eta = c_1 t^\mu + d_1 t^{\mu+1} + \cdots \quad (c \neq 0, c_1 \neq 0).$$

The first equation can be solved for t as a series in $\tau = (\xi - \alpha)^{1/\rho}$ defining ρ distinct values of t for each ξ near α when the ρ values of τ are substituted. When this series is substituted in the second equation the result has the form

$$(10)\qquad \eta = \beta\tau^\mu + \beta'\tau^{\mu'} + \cdots.$$

The q values $\eta_1, \cdots, \eta_q$ of $\eta(x, y)$ corresponding to a value of ξ near $\xi = \alpha$ are defined by a number of equations (10), one for each of the zeros of $\xi(x, y) - \alpha$. The sum of the orders ρ of these is q. For values ξ near $\xi = \infty$ the only difference is that the parameter in (10) is $\tau = (1/\xi)^{1/\rho}$.

The symmetric functions of the values $\eta_1, \cdots, \eta_q$ corresponding to a given ξ are rational functions of ξ. For they are clearly representable by series in fractional powers of $(\xi - \alpha)$ or $1/\xi$ near $\xi = \alpha$ or $\xi = \infty$, and these series can have only integral exponents since the symmetric functions which they represent are single-valued in ξ. Hence the symmetric functions, being single-valued in ξ and having no singularities except poles, are rational in ξ. It follows now that the product $(\eta - \eta_1)(\eta - \eta_2) \cdots (\eta - \eta_q)$ is a polynomial of degree q in η with coefficients rational in ξ, and when cleared of fractions it becomes a polynomial $\phi(\xi, \eta)$ in both ξ and η. The equation $\phi(\xi, \eta) = 0$ of degree q in η is satisfied by all of the values of $\xi(x, y)$, $\eta(x, y)$ at places on T.

The polynomial $\phi(\xi, \eta)$ is either irreducible or else the power of a single irreducible factor. In fact, each irreducible factor of $\phi(\xi, \eta)$, being made zero by one at least of the expansions (10), must also be caused to vanish identically by (9) and all of its continuations on the surface T. But in that case the irreducible factor must be satisfied by all of the roots $\eta_1, \cdots, \eta_q$, and the factors of $\phi(\xi, \eta)$ must all be identical. The important case when $\phi(\xi, \eta)$ itself is irreducible will evidently occur if and only if there is at least one value ξ for which all of the values $\eta_1, \cdots, \eta_q$ are distinct.

When $\phi(\xi, \eta)$ is itself irreducible the equation $\phi(\xi, \eta) = 0$ is birationally related to $f(x, y) = 0$. For then the set of series (10) corresponding to $\xi = \alpha$ must define distinct values $\eta_1, \cdots, \eta_q$ for values of ξ near $\xi = \alpha$, and the series themselves must all be distinct. To each one of them corresponds therefore a unique place on the surface T, which is equivalent to saying that there corresponds to each place on the Riemann surface U of $\phi(\xi, \eta) = 0$ a unique place P on the surface T. The values x, y belonging to P are therefore single-valued functions on the surface U. They have at most poles, since as

functions of t they have only poles, and since t is expressible in positive integral powers of τ by solving the equation

$$\tau = t(c + dt + \cdots)^{1/\rho}$$

obtained from the first equation (9) and the equation $\xi - \alpha = \tau^\rho$. Under these circumstances x and y are necessarily rational functions of ξ, η, and the curves $f(x, y) = 0$ and $\phi(\xi, \eta) = 0$ are birationally related.

4. **Transformation to a curve with ordinary double points only in the projective plane.** In this section it is proposed to prove the following principal theorem of this paper:

An irreducible algebraic curve $f(x, y) = 0$ can always be transformed by a birational transformation into a second such curve $\phi(\xi, \eta) = 0$ having no singularities in the projective $\xi\eta$-plane except ordinary double points.

The method of proof is to construct, with the help of the functions $\sigma_1, \cdots, \sigma_\nu$ of Section 2, a pair of rational functions $\xi(x, y), \eta(x, y)$ satisfying an algebraic equation $\phi(\xi, \eta) = 0$ birationally related to $f(x, y) = 0$ and having the properties 1 and 2 of Lemma 1.

Consider now a divisor $\mathfrak{D} = P_1^{-1} \cdots P_q^{-1}$ with $q > 2p + 2$, as in Section 2, having the linearly independent multiples $\sigma_1, \cdots, \sigma_\nu$. A particular function

$$\xi(x, y) = u_1\sigma_1 + \cdots + u_\nu\sigma_\nu$$

can be chosen with the constants $u_1, \cdots, u_\nu$ fixed once for all so that at each of the places $P_1, \cdots, P_q$ the order of ξ is exactly -1 (Lemma 2, Corollary 1 for $k = 1$). For a second function of the family

$$\eta(x, y) = v_1\sigma_1 + \cdots + v_\nu\sigma_\nu \tag{11}$$

the expansions (10) corresponding to $\xi = \infty$ arise at the places $P_1, \cdots, P_q$, and are readily seen to have the form

$$\eta = \xi\left(\beta_k + \beta_k'\frac{1}{\xi} + \cdots\right) \qquad (k = 1, \cdots, q). \tag{12}$$

The constants β_k will be distinct provided that $v_1, \cdots, v_\nu$ do not satisfy a certain system L of linear equations (Lemma 2, Corollary 1 for $k = 2$). Under these circumstances the values $\eta_1, \cdots, \eta_q$ corresponding to a value of ξ near $\xi = \infty$ will be distinct, and ξ and η will satisfy an irreducible algebraic equation $\phi(\xi, \eta) = 0$ of degree q in η birationally related to $f(x, y) = 0$. Furthermore, since η is infinite only at the places $P_1, \cdots, P_q$ where ξ is also infinite, the only poles of the function $\eta(\xi)$ defined by $\phi = 0$ are those given by the q expansions (12), and these have now the properties prescribed in 1 of Lemma 1.

The discriminant of η with respect to ξ,

$$D(v_1, \cdots, v_\nu, \xi) = \prod_{i<j}(\eta_i - \eta_j)^2 \qquad (i, j = 1, \cdots, q), \tag{13}$$

is a polynomial in $v_1, \cdots, v_\nu, \xi$. For it is symmetric in $\eta_1, \cdots, \eta_q$ and therefore rational in ξ, and it must be a polynomial in ξ since it is finite near every finite value of ξ. It may be represented as a product $W(v_1, \cdots, v_\nu, \xi)\Delta(\xi)$ where $\Delta(\xi)$ contains all factors of D containing only ξ, and where W has consequently no factor in ξ alone.

Near a finite value $\xi = \alpha$ the q values of η are defined by a number of series (10) which now have the form

$$\eta = \beta + \beta'(\xi - \alpha)^{1/\rho} + \cdots \tag{14}$$

since η has no poles except at the places $P_1, \cdots, P_q$ where ξ is also infinite. The ρ values of η defined by one of these equations provide $\rho(\rho - 1)/2$ factors $\eta_i - \eta_j$ in D, each repeated twice, and each of order at least $1/\rho$ in $(\xi - \alpha)$. Hence they contribute at least the factor $(\xi - \alpha)^{\rho-1}$ to $\Delta(\xi)$. If the coefficients $v_1, \cdots, v_\nu$ in (11) do not satisfy a certain system of linear equations, then the coefficients β' in the expansions (14) corresponding to $\xi = \alpha$ will all be different from zero (Lemma 2, Corollary 2), and the constants β will all be distinct (Lemma 2, Corollary 1 for $k = 2$). For a suitable special choice of the coefficients v, therefore, the discriminant D will contain no more powers of $(\xi - \alpha)$ than the product of the factors $(\xi - \alpha)^{\rho-1}$ corresponding to the different expansions (14) for $\xi = \alpha$, and it can consequently contain no more when the v's are indeterminates. It follows readily that $\Delta(\xi)$ is exactly the product $\Pi(\xi - \alpha)^{\rho-1}$ taken for all the expansions at branch values $\xi = \alpha$, i.e., the values for which the function $\xi(x, y) - \alpha$ has multiple roots.

The factor $W(v_1, \cdots, v_\nu, \xi)$ of the discriminant is irreducible or else a product of a number of irreducible factors which are also polynomials in $v_1, \cdots, v_\nu, \xi$. Near a value $\xi = \xi_0$ which is not a branch value of ξ the discriminant D, thought of as a polynomial in $v_1, \cdots, v_\nu$, is completely decomposable into a product of factors

$$\eta_i - \eta_j = v_1(\sigma_{1i} - \sigma_{1j}) + \cdots + v_\nu(\sigma_{\nu i} - \sigma_{\nu j}) \tag{15}$$

linear in $v_1, \cdots, v_\nu$ and with coefficients power series in $\xi - \xi_0$ with constant terms not all zero (Lemma 2, Corollary 1 for $k = 2$). Each irreducible polynomial factor of D must therefore be expressible near $\xi = \xi_0$ as the product of a number of distinct factors of the form (15), and each must occur at least twice in W since each linear factor (15) occurs twice in D. An irreducible factor of W could not occur more than twice, since otherwise there would have to be at least two linear factors $\eta_i - \eta_j$, $\eta_k - \eta_l$ having $(i, j) \neq (k, l)$ and with proportional coefficients. This would imply a relation

$$\eta_i - \eta_j = (\eta_k - \eta_l)g$$

holding identically in $v_1, \cdots, v_\nu$, where g is a series in $(\xi - \xi_0)$ with constant

term different from zero. But this would contradict the validity of Lemma 2 at $\xi = \xi_0$, or its Corollary 1 for $k = 3$ when only three of the integers i, j, k, l are distinct. It is clear then that each irreducible factor in W occurs exactly twice, and D has the form

$$D(v_1, \cdots, v_r, \xi) = V(v_1, \cdots, v_r, \xi)^2 \Delta(\xi),$$

where V is a product of distinct irreducible polynomials in $v_1, \cdots, v_r, \xi$ containing no factor in ξ alone.

Suppose now that the coefficients $v_1, \cdots, v_r$ have numerical values not satisfying the linear system of equations L. Then, as has been shown above, the functions $\xi(x, y)$, $\eta(x, y)$ satisfy an irreducible algebraic equation $\phi(\xi, \eta) = 0$ birationally related to $f(x, y) = 0$, and the only poles of the function $\eta(\xi)$ defined by $\phi(\xi, \eta) = 0$ are given by the expansions (12) which have the form prescribed in 1 of Lemma 1. If furthermore $v_1, \cdots, v_r$ are so chosen that the roots of the irreducible factors of $V(v_1, \cdots, v_r, \xi)$ are distinct from each other and from those of $\Delta(\xi)$, then D has the form prescribed in 2 of Lemma 1. Hence $\phi(\xi, \eta) = 0$ has in the projective plane only double points with distinct tangents.

5. **Transformation to a curve with ordinary double points only in the function-theoretic plane.** The projective xy-plane is the extension of the euclidean xy-plane which is obtained by introducing the homogeneous coördinates $x = x_1/x_3$, $y = x_2/x_3$. The function-theoretic plane, on the other hand, is obtained when the homogeneous coördinates $x = x_1/x_2$, $y = y_1/y_2$ are used. In the former case the infinite region is the line $x_3 = 0$. In the latter case it is the totality of points whose non-homogeneous coördinates have the form (∞, y) or (x, ∞), or whose homogeneous coördinates have $x_2 = 0$ or $y_2 = 0$.

An irreducible algebraic curve $f(x, y) = 0$ is said to have ordinary double points only in the function-theoretic xy-plane if its singularities in the finite part of the plane are ordinary double points, and if its points at infinity provide only simple points or double points with distinct tangents when they are transformed into the finite part of the plane by means of one or both of the transformations $x = 1/x'$ or $y = 1/y'$. It is possible to transform an irreducible algebraic curve $f(x, y) = 0$ into one with only ordinary double points in the function-theoretic plane by a birational transformation of the form

$$\xi = x, \qquad \eta = \eta(x, y),$$

as will be shown in the following paragraphs. The first step in the proof of this statement is the following lemma analogous to Lemma 1 of Section 1.

LEMMA 3. *An irreducible algebraic curve $f(x, y) = 0$ of degree n in y has no singularities except double points with distinct tangents in the function-theoretic*

plane, provided that the function $y(x)$ defined by the equation $f = 0$ has the following properties:

1. *at $x = \infty$ the function $y(x)$ has n distinct finite values;*

2. *the poles of $y(x)$ are simple and at values $x = a$ distinct from each other and from the branch values of x;*

3. *the discriminant of $y(x)$ has the property 3 of Lemma 1.*

It is clear from 1 that the points (∞, y) on the curve are simple when transformed into finite points by the transformation $x = 1/x'$, and from 2 the same is true of the points (x, ∞) when transformed by $y = 1/y'$.

The finite roots of $f(x_0, y)$ for a value $x = x_0$ over which $y(x)$ has a pole are all simple, by 2, and hence they provide only simple points of the curve

$$f(x, y) = 0.$$

The discriminant of $y(x)$ is by definition

$$D(x) = A^{2n-2} \prod_{i<j} (y_i - y_j)^2 \qquad (i, j = 1, \cdots, n),$$

where A is the product of the linear factors $x - x_0$ belonging to values x_0 for which $y(x)$ has poles. By a repetition of the argument of Section 1 it is provable that $f(x, y) = 0$ has only double points with distinct tangents corresponding to values $x = a$ other than those discussed above, and the lemma is therefore proved.

To construct a rational function $\eta(x, y)$ whose irreducible algebraic equation $\phi(x, \eta) = 0$ is birationally related to $f(x, y) = 0$ and has the properties 1, 2, 3 of Lemma 3, one may start from a divisor $\mathfrak{D} = P_1^{-1} \cdots P_q^{-1}$ with $q > 2p + 2$, as in Section 2, whose places P_i on T correspond to finite values of x distinct from each other and from the branch values. If the coefficients $v_1, \cdots, v_\nu$ in the expression

$$\eta(x, y) = v_1 \sigma_1 + \cdots + v_\nu \sigma_\nu$$

do not satisfy a certain system L_1 of linear equations, the function $\eta(x, y)$ will have distinct finite values at $x = \infty$ (Lemma 2, Corollary 1 for $k = 2$) and a pole of order one at each place P_i (Lemma 2, Corollary 1 for $k = 1$). Since under these circumstances the n values of η corresponding to a value x near $x = \infty$ are all distinct, it follows that $\eta(x, y)$ will satisfy an algebraic equation $\phi(x, \eta) = 0$ birationally related to $f(x, y) = 0$ and having the properties 1 and 2 of Lemma 3.

The discriminant of $\eta(x, y)$ with respect to x is by definition

$$D(v_1, \cdots, v_\nu, \xi) = A^{2n-2} \prod_{i<j} (\eta_i - \eta_j)^2 \quad (i, j = 1, \cdots, n),$$

where A is the product of the factors $(x - x_0)$ belonging to the places P_i of D, and where $\eta_1, \cdots, \eta_n$ are the n values of $\eta(x, y)$ for a given value x. This expression for D is symmetric in $\eta_1, \cdots, \eta_n$ and therefore rational in x. It can become infinite only at the values x_0 corresponding to the places P_i. At such a value x_0 only one of the values η_i has a simple pole. There are therefore only $n - 1$ of the factors $\eta_1 - \eta_j$ which have simple poles, and in D these are annulled by the factor A^{2n-2}. Hence D is finite for all finite values of x and is a polynomial in $v_1, \cdots, v_r, x$. It is expressible in the form

$$W(v_1, \cdots, v_r, x)\Delta(x)$$

where W contains no factor in x alone.

It may now be proved, exactly as in Section 4, that $\Delta(x)$ is the product $\Pi(x - a)^{r-1}$ formed for all the branch places on T, and that W is the square of a product $V(v_1, \cdots, v_r, x)$ of distinct irreducible factors. Hence if $v_1, \cdots, v_r$ are chosen not satisfying the system L_1 of linear equations, and so that the roots of V are distinct from each other and those of $\Delta(x)$, then $\phi(x, \eta) = 0$ will have the properties 1, 2, and 3 of Lemma 3, and the curve which it defines in the function-theoretic plane will have only double points with distinct tangents.

6. **Homogeneous coördinates.** The branches of an irreducible algebraic curve $f(x, y) = 0$ have for $x = \infty$ and $x = a$ the forms

$$x = t^{-r}, \qquad y = bt^{\mu} + b't^{\mu'} + \cdots,$$
$$x = a + t^{r}, \qquad y = bt^{\mu} + b't^{\mu'} + \cdots.$$

These, with their transforms after replacing t by a series $ct + dt^2 + \cdots$ $(c \neq 0)$, are included in the type

$$x = P(t), \qquad y = Q(t), \tag{16}$$

where P and Q are series in integral powers of t having at most a finite number of terms with negative exponents.

In terms of homogeneous coördinates $x = x_1/x_3$, $y = x_2/x_3$ the branch (16) takes the form

$$x_l = a_l + \alpha_l t + \cdots \qquad (l = 1, 2, 3) \tag{17}$$

where $(a_1, a_2, a_3) \neq (0, 0, 0)$. A branch of this character is said to be linear if the matrix

$$\begin{vmatrix} a_1 & a_2 & a_3 \\ \alpha_1 & \alpha_2 & \alpha_3 \end{vmatrix}$$

is of rank 2. Two such branches with the same center $a_1 : a_2 : a_3$ have distinct tangents if the determinant

$$\begin{vmatrix} a_1 & a_2 & a_3 \\ \alpha_1 & \alpha_2 & \alpha_3 \\ \alpha_1' & \alpha_2' & \alpha_3' \end{vmatrix}$$

is different from zero, where α_1', α_2', α_3' are the coefficients of t in the second branch.

In the projective plane a curve $f(x, y) = 0$ has no singularities except double points with distinct tangents if all of its branches are linear, if the same center is never shared by more than two branches, and if every pair of branches with a common center has distinct tangents. It is readily provable that the branches (17) of a curve $f(x, y) = 0$ such as is described in Lemma 1 of Section 1 have these properties, and that the properties themselves are invariant under transformations of the parameter t and projective transformations

$$x_i = \alpha_{i1}\,\xi_1 + \alpha_{i2}\,\xi_2 + \alpha_{i3}\,\xi_3 \qquad (i = 1, 2, 3).$$

In terms of homogeneous coördinates $x = x_1/x_2$, $y = y_1/y_2$ the branches (16) take the form

$$x_i = a_i + \alpha_i t + \cdots,$$
$$y_i = b_i + \beta_i t + \cdots \qquad (i = 1, 2).$$

Such a branch is said to be linear if the determinants

$$\begin{vmatrix} a_1 & \alpha_1 \\ a_2 & \alpha_2 \end{vmatrix}, \qquad \begin{vmatrix} b_1 & \beta_1 \\ b_2 & \beta_2 \end{vmatrix}$$

are not both zero. Two branches with the same center

$$(a_1 : a_2 | b_1 : b_2) = (a_1' : a_2' | b_1' : b_2')$$

have distinct tangents if the expression

$$\begin{vmatrix} a_1 & a_1' \\ a_2 & a_2' \end{vmatrix} \begin{vmatrix} b_1' & \beta_1 \\ b_2' & \beta_2 \end{vmatrix} - \begin{vmatrix} a_1' & \alpha_1 \\ a_2' & \alpha_2 \end{vmatrix} \begin{vmatrix} b_1 & \beta_1' \\ b_2 & \beta_2' \end{vmatrix}$$

is different from zero, where the primes designate the coefficients of the second branch.

A curve with no singularities except double points with distinct tangents in the function-theoretic plane is now defined as in the next to last paragraph above. The curves described in Lemma 3 of Section 5 have these properties, and the properties themselves are invariant under transformations of the parameter t and the transformations

$$x_i = \alpha_{i1}\,\xi_1 + \alpha_{i2}\,\xi_2,$$
$$y_i = \beta_{i1}\,\eta_1 + \beta_{i2}\,\eta_2 \qquad (i = 1, 2).$$

THE UNIVERSITY OF CHICAGO,
CHICAGO, ILL.

A SYMBOLIC THEORY OF FORMAL MODULAR COVARIANTS*

BY

OLIVE C. HAZLETT

PART I. INTRODUCTION

1. **Prologue.** Thus far, very little has been published on the general theory of formal modular invariants or covariants. Workers have, on the whole, obtained results for special, more or less isolated, cases; and although some beautiful and important general theorems have been proved, they are more or less unrelated. This is, of course, only natural in any division of knowledge in its formative state.

Nevertheless, no worker in the field could fail to be conscious of a certain uniformity common to the special cases that have been studied in detail; though (alas!) this uniformity usually appeared to be broken ruthlessly in the next case studied. This breaking of an apparent law signified, however, merely that we did not know these special cases with a sufficient thoroughness of illuminating detail, or were trying unwittingly to make the laws conform to certain standards, unconsciously preconceived. This latter handicap was laid on us naturally enough by our thorough knowledge of algebraic invariants and the fact that this newer kind of covariants is, in many ways, strikingly like the older, classic covariants, though so tantalisingly different.

Their similarity and their difference show themselves in the very beginning of the study: in the definitions, in the simplest examples. Perhaps the differences that first come to mind are those which are inherent in the fields of definition, which, in the case of classic covariants, is the field of reals or ordinary complex numbers and, in the case of modular covariants, is a Galois field, $GF[p^n]$, of order p^n. These differences are too obvious to mention in detail, but one who has studied the beautiful proofs given by the old masters of invariant theory has been forced to the conclusion that most of the proofs seemed to use the properties of a field of characteristic zero, not in some accidental manner, but rather in veriest necessity.

Growing from the surface differences between the two fields are two very important distinguishing characteristics of the two kinds of covariants. It

* Part II was presented to the Society, September 7, 1920; Part III, December 28, 1921; Parts IV and V, December 27, 1922.

The reading of the literature in connection with this paper was much facilitated by the purchase of books with a grant made by the American Association for the Advancement of Science and this help is herewith gratefully acknowledged.

is well known that an algebraic covariant is necessarily such that all of its terms have the same weight—i.e., it is isobaric; but a modular covariant is not necessarily isobaric. Nevertheless, although its terms are not in general of exactly the same weight, their weights can differ at most by multiples of $p^n - 1$. Moreover, if two classic invariants are identical in value for all marks of the field, they are necessarily identical in form, and conversely; whereas, if two formal modular invariants are identical for all marks of the field of definition, they are not necessarily identical in form. Nevertheless, although two such invariants can be of different degree and appear quite different in form at a casual glance, yet they necessarily have in common certain fundamental characteristics.* There are other differences that will occur to any worker in the field, but I think that I have mentioned the most refractory.

As indicated above, in spite of the great differences between the algebraic covariants and formal modular covariants, there are certain fundamental likenesses which are more easily sensed than they are analyzed. A feeling that there is some theory which underlies all the special cases, and yet which is comparatively simple, made the writer try to crystallize this theory into words.

In the spring of 1918, came the feeling that the theory of formal modular covariants of a binary form, f, for the field $GF[p^n]$, must be, in essentials, equivalent to the theory of simultaneous algebraic covariants of f and certain other forms obtained from f by replacing the coefficients of f by their respective p^nth powers. This is natural enough, since these powers of the coefficients of f are cogredient with the coefficients of f for the transformations of the group used, and this is the only way in which a formal modular covariant differs from an algebraic covariant of f. Then there appeared other indications that there is an intimate relation between formal modular covariants of f and algebraic covariants of a system of forms consisting of f and related forms. This paper is an attempt to put in systematic form the theorems which emerged when these eventually crystallized.

2. **Summary of literature.** Let

$$f(a; x) = a_0 x^m + a_1 x^{m-1} x_2 + \cdots + a_{m-1} x_1 x^{m-1} + a_m x^m$$

be a binary form and let G be a group of linear transformations

$$x_i = \xi_i x_1' + \eta_i x_2' \qquad (i = 1, 2) \tag{1}$$

whose coefficients, the ξ's and η's, are in the field F. If $I(a)$ be a polynomial in the a's which, under all transformations of the group, is transformed into $I(a')$ such that

$$I(a') = D^w I(a) \quad \text{(in the field)} \tag{2}$$

* Some of these characteristics are indicated in a paper by Hazlett, these Transactions, vol. 22 (1921), pp. 144–157, especially p. 145.

where

$$D = (\xi\eta) = \xi_1 \eta_2 - \xi_2 \eta_1,$$

then I is said to be an invariant of f under the group G. Similarly, if $C(a; x)$ is a polynomial in the a's and the x's which, under all transformations of the group G, is transformed into $C(a'; x')$ such that

$$C(a'; x') = D^w C(a; x) \quad \text{(in the field)}, \tag{3}$$

then C is said to be a covariant of f under the group G. Also, we speak of invariants and covariants of a system S of binary forms f_i.

When the field of definition, F, is the field of all reals or the field of all ordinary complex numbers, the covariants of S are the ordinary covariants of S of the classic invariant theory, with which are associated the names of Cayley, Sylvester, Hermite, Aronhold, Gordan, Clebsch and Hilbert.

When the field of definition, F, is the Galois field, $GF[p^n]$, of order p^n, the covariants of S are called modular covariants. Here the ξ's and η's are marks of the Galois field, $GF[p^n]$, of order p^n* defined by the prime p and an algebraic equation, $P(X) = 0$, of degree n; and thus (2) and (3) are congruences, reduced modulis p and $P(X)$. This means that, in (2) and (3), the left member is identical with the right member if we replace the p^nth power of ξ_i and of η_i by ξ_i and η_i respectively, in view of Galois' generalization of Fermat's theorem.† For a modular field there are two different types of covariants. If the coefficients of the forms (the a's) range over the marks of $GF[p^n]$, and if (2) and (3) are congruences which are true if Fermat's theorem is applied, not only to the ξ's and η's, but also to the a's, then the covariants are called simply *modular covariants*. If, on the other hand, the a's are independent variables, Fermat's theorem does not apply to them; and hence, in (2) and (3), the left member is understood to be identical with the right member without any reduction in the exponents of the a's. Such covariants are called *formal*

* Let p be any prime and let $P(X) = 0$ be any algebraic equation of degree n which has its coefficients integers reduced modulo p and which is irreducible, modulo p. Then, if we reduce any polynomial in X modulis p and $P(X)$, we obtain a polynomial of the form

$$M(X) = c_{n-1} X^{n-1} + c_{n-2} X^{n-2} + \cdots + c_1 X + c_0$$

where each of the c's is an integer of the set $0, 1, \cdots, p-1$. The totality of all polynomials congruent modulis p and $P(X)$ to the same $M(X)$ is said to form a class of residues. Since there is a class of residues for each set of values for the c's, there are p^n such classes. The totality of the classes of residues are closed under addition, subtraction, multiplication and division (provided the divisor is not zero) and so constitute a field containing p^n marks or elements. Any two such fields are the same for a given p and a given n. This is called a Galois field of order p^n and is denoted by $GF[p^n]$. Moreover, any finite field is simply isomorphic with a Galois field.

† If a is any mark of $GF[p^n]$, defined by p and $P(X)$, then $a^{p^n} \equiv a \pmod{p, P(X)}$. See any standard work on Galois fields, such as Dickson, *Linear Groups*, Teubner, 1901, p. 11; Serret, *Cours d'Algèbre Supérieure*, vol. 2, p. 180.

modular covariants. If the covariant is independent of the variables, it is called a *modular invariant* or a *formal modular invariant* according as the a's are marks of the field or independent variables.

Although Hurwitz introduced the notion of formal modular invariants in connection with the determination of the number of solutions of higher congruences,* Dickson discovered them independently four years later from a different point of view.† Practically all important results in the theory are due to Dickson.‡

Clearly, every formal modular covariant of a system S is a modular covariant of S, though not every modular covariant of S is also a formal modular covariant of S. For a modular invariant, I, of S is concerned solely with the values of I for the different sets of values of the a's in the field of definition, whereas a formal modular invariant of S is concerned not only with the values but also with the form of I. This statement has to be modified somewhat for covariants that are not invariants, since, in (2) and (3), the left member is the same as the right member in a purely formal sense as far as the x's are concerned, for both kinds of covariants. For example, if a is any one of the coefficients of f, then $a^{p^n} - a$ always has the value zero when a is a mark of $GF[p^n]$ and hence it is a modular invariant of f; but it is not a formal modular invariant, since it is changed in form if we interchange x_1 and x_2. To take a less trivial example,

$$q = (a + c)(b^2 + ac - 1)$$

is a modular invariant of $f = ax_1^2 + 2bx_1x_2 + cx_2^2$, mod 3, but is not a formal invariant, since $a + c$ goes into $-a - b + c$ under the transformation $x_1 = x_1' + x_2'$, $x_2 = x_2'$.

In 1909,§ Dickson studied modular invariants from a different point of view and introduced the notion of classes of forms. This enables one to see to the very heart of the theory and the finiteness theorem follows almost directly from the definition. Later,‖ he proved the finiteness theorem for modular covariants.

* *Ueber höhere Kongruenzen*, Archiv der Mathematik und Physik, ser. 3, vol. 5 (1903), pp. 17–27.

† *Invariants of binary forms under modular transformations*, these Transactions, vol. 8 (1907), pp. 205–232.

‡ Anyone wishing to gain familiarity with this beautiful theory should read the article by Hurwitz mentioned above and Dickson's papers, of which the most fundamental are I. *General theory of modular invariants*, these Transactions, vol. 10 (1909), pp. 123–158; II. *Proof of the finiteness of modular covariants*, ibid., vol. 14 (1913), pp. 299–310. Also one should read the brief but important paper by Miss Sanderson, *Formal modular invariants with application to binary modular covariants*, these Transactions, vol. 14 (1913), pp. 489–500. All results obtained up to 1914 are summarized in Dickson's *Madison Colloquium Lectures*, and all essential results published up to August, 1922, are summarized briefly in Chapter 19 of his *History of the Theory of Numbers*, vol. 3.

§ Dickson, I.

‖ Dickson, II.

There is, however, an intimate relation between modular invariants and formal modular invariants. For Miss Sanderson's theorem* tells us that, corresponding to any modular invariant, i, of a system S under any group G of linear transformations with coefficients in $GF[p^n]$, there is a formal invariant, I, under G such that $I \equiv i$ for all sets of values, in the field, of the coefficients of the system S. This enabled her to construct modular covariants of a system S of binary forms, f_i, from the modular invariants of another system S', consisting of the forms f_i and an additional linear form with coefficients x_2 and $-x_1$. Moreover, every modular covariant of S is a polynomial in the universal covariant, $L = x_1^{p^n} x_2 - x_1 x_2^{p^n}$, and the modular covariants obtained in the manner indicated by Miss Sanderson's theorem.† Since every algebraic covariant of S is obtained from the algebraic invariants of S' without the need of any additional covariant, analogous to L, this result shows one of the fundamental differences between algebraic and modular covariants.

In connection with this last remark, several minor results are of interest, as they point the way toward more general results. In 1920, it was shown that, for the field $GF[p^n]$ with $p \neq 2$, every modular covariant of a binary form, f, whose degree is not divisible by p, is expressible as a rational function of the universal modular covariants, L and Q, and of algebraic covariants of f. For formal modular covariants of f, there is a theorem which is more complicated in statement but similar in essence.‡ Then, in 1922, W. L. G. Williams announced that every formal modular seminvariant of f (aside from a power of a_0) is a polynomial in the algebraic protomorphs and in

$$\beta = a_1^p - a_1 a_0^{p-1}$$

for $GF[p]$ when the form f is such that binomial coefficients can be used.§ He also proved the analogous theorem for formal seminvariants of two or more binary forms. In this latter, it is interesting to note that, to a fundamental set of algebraic protomorphs, it is necessary to adjoin only one new seminvariant, namely, one of the same type as β but formed for any one of the forms f_i. It will be seen that these results indicate relations between modular covariants and algebraic covariants, in spite of their superficial differences.

3. **Summary of this paper.** We first explain a symbolic notation for formal modular covariants which is like that generally used for algebraic covariants in most respects, but is (necessarily) different from the latter in one essential.

* These Transactions, vol. 14 (1913), pp. 489–500.

† Hazlett, *A theorem on modular covariants*, these Transactions, vol. 21 (1920), pp. 247–254.

‡ Hazlett, *Associated forms in the general theory of modular covariants*, American Journal of Mathematics, vol. 43 (1921), pp. 194–196.

§ *Fundamental systems of protomorphic formal modular seminvariants of binary forms*, read before the Society (Rochester, N. Y.), Sept. 8, 1922. I saw this article in MS. after announcing to the Society (Toronto), on December 28, 1921, the results of Part III of this paper.

This notation, however, has the advantage that it can be used for both algebraic covariants and formal modular covariants; and thus, by its use, the proof of every theorem in this paper is such that it applies both to the modular and the non-modular cases. To help the reader appreciate the diversity of these two interpretations of this theory, the more important theorems are applied to a number of special cases.

As in the symbolic theory of algebraic covariants, it is readily seen that all formal modular covariants are polynomials in a finite number of symbolic expressions having the invariant property. Then are proved several other elementary theorems which assume a simple and familiar form for algebraic covariants, but which, for formal modular covariants, assume a form rather strikingly different.

It is then shown that every isobaric formal modular covariant, C, of the system S is congruent, in the field, to a function of the coefficients of S which is (in a certain general sense) an algebraic covariant of S. When, however, C is not isobaric, it is not congruent to an algebraic covariant of S; but it is, nevertheless, congruent to an algebraic covariant of an enlarged system, S', consisting of the forms of S and other forms obtained from the forms of S by replacing the coefficients of S by their p^nth powers. By various devices, there are expressed in symbolic form the formal modular invariants of a fundamental set for the cubic, modulo 2, and for the quadratic, modulo 3; and the above theorems are verified.

Finally, by using the same theorems, we prove that the set of all formal modular covariants of any binary system, S, with respect to the Galois field, $GF[p^n]$, is such that (1) all syzygies among them are consequences of a finite number of such syzygies; and (2) all formal modular covariants are expressible as polynomials in a finite number of such covariants.

Part II. Symbolic notation

4. **Explanation of the notation.** As in the theory of algebraic covariants, formal modular covariants of a binary form, f, assume a form which is both simple and elegant when the form f is expressed as a product of symbolic linear factors. In the theory of algebraic covariants, it is customary to express a binary form f of degree m as a symbolic mth power; but, for reasons which were explained elsewhere in detail,* it is not possible to express the general form f of degree m as a symbolic mth power and then express all formal modular covariants by means of these symbols for the general Galois field.

For this reason, represent

$$f = a_0 x_1^m + a_1 x_1^{m-1} x_2 + \cdots$$

* *New proofs of certain finiteness theorems in the theory of modular covariants*, these Transactions, vol. 22 (1921), pp. 152–153.

as

$$f = \pi\alpha_x = \pi(\alpha_1 x_1 + \alpha_2 x_2) = (\alpha_1 x_1 + \alpha_2 x_2)(\beta_1 x_1 + \beta_2 x_2)\cdots$$

where it is understood that there are m symbolic factors which are symbolically distinct, and consider its formal modular covariants under the group of transformations (1) where the ξ's and η's are any marks of the Galois field, $GF[p^n]$, of order p^n such that the determinant is not zero in the field. Then, as has been proved elsewhere,* every rational integral covariant (both classic and modular) of f is a polynomial in the symbols $\alpha, \beta, \cdots$ and in the x's. The converse is also true; though, in order that a polynomial in the x's and in the symbols be rational in the a's, (i) it must be symmetric in the m pairs (α_1, α_2), $(\beta_1, \beta_2), \cdots$ and such that each term contains as many α's as β's, as many β's as γ's, etc.; or (ii) it must be a sum of a number of expressions described under (i). Similar remarks apply to covariants of several binary forms. For convenience, any polynomial in the x's and the symbols which has the invariantive property under a transformation of the group will be called a symbolic covariant, even if it do not satisfy the conditions that it be rational in the a's.

5. **Fundamental set of symbolic covariants.** Early in the symbolic theory of algebraic covariants, it is proved that all algebraic invariants are expressible as polynomials in a finite number of symbolic invariants of the type

$$(\alpha\beta) = \alpha_1\beta_2 - \alpha_2\beta_1.$$

Similarly, all algebraic covariants are expressible as polynomials in a finite number of symbolic covariants of the type $(\alpha\beta)$ and $\alpha_x = \alpha_1 x_1 + \alpha_2 x_2$. For formal modular covariants we have a similar theorem, though we have to use new types of symbolic covariants.

For simplicity, we shall first consider covariants of a single form, $f = \pi\alpha_x$. Now it is known that the pair (α_1, α_2) is pseudo-cogredient with $(x_2, -x_1)$.† Accordingly, any formal modular covariant may be regarded, for any purpose not involving the notions of weight or index, as if it were actually an invariant of $m+1$ cogredient pairs, consisting of the pairs of symbols and the pair $(x_2, -x_1)$. But it is known that the formal modular invariants of a number of cogredient points have the finiteness property.‡ Notice that this argument applies equally well to a system of binary forms, and thus we have

THEOREM I. *All rational integral formal modular covariants of a system of binary forms,* $f_i = \pi(\alpha_1 x_1 + \alpha_2 x_2) = \alpha_x\beta_x\cdots$, *with respect to the Galois*

* Preceding reference, top of p. 153.

† That is, the two pairs are cogredient aside from a multiplicative factor which is a power of the determinant of the transformation.

‡ F. B. Wiley, *Proof of the finiteness of the modular covariants of a system of binary forms and cogredient points*, these Transactions, vol. 15 (1914), pp. 431–438.

field $GF[p^n]$, of order p^n, are polynomials in a finite number of symbolic covariants which are polynomials in the x's and in the α's, β's, $\cdots$.

6. **Illustration of Theorem I.** Dickson* has shown that all formal modular invariants of the binary quadratic $f = a_0 x_1^2 + 2a_1 x_1 x_2 + a_2 x_2^2$, modulo 3, are polynomials in

$$\Delta = a_1^2 - a_0 a_2, \qquad J = a_0 \gamma_0 = a_0 (a_0 + a_1 + a_2)(a_0 + 2a_1 + a_2) a_2,$$
$$B = \beta\gamma_1 \equiv a_1 (a_1^2 - a_0^2)(a_2 - a_0)(a_2^2 - a_1^2),$$
$$\Gamma = (a_0 + a_2)(2a_0 + 2a_1 + a_2)(2a_0 + a_1 + a_2).$$

If we express f in symbolic form (§ 4) as $f = (\alpha_1 x_1 + \alpha_2 x_2)(\beta_1 x_1 + \beta_2 x_2)$, then $a_0 = \alpha_1\beta_1$, $a_1 \equiv -(\alpha_1\beta_2 + \alpha_2\beta_1) \mod 3$, $a_2 = \alpha_2\beta_2$. Hence it is readily seen that

$$\Delta \equiv (\alpha\beta)^2, \qquad J = (\alpha^3\alpha)(\beta^3\beta),$$
$$B \equiv \sum (\alpha^3\beta)^2 (\beta^3\beta),$$
$$\Gamma \equiv -N_1 \qquad (\mod 3),$$

where

$$N_1 = [(\alpha^3\beta)^2 (\beta^3\beta) - (\beta^3\alpha)^2 (\alpha^2\alpha)]/(\alpha\beta)^3$$

is one of a fundamental set of formal invariants of the pairs (α_1, α_2) and (β_1, β_2) and is an integral function of the α's and β's. Thus Theorem I is verified for this case. Note, also, that the rational invariants of f are expressible as rational functions in the determinantal symbolic invariants, $(\alpha\beta)$, $(\alpha^3\beta)$, $(\alpha\beta^3)$, $(\alpha^3\alpha)$, $(\beta^3\beta)$.

Part III. Relation between classic and modular covariants

7. **Two kinds of congruences.** In the following sections, we shall frequently have to distinguish between two kinds of congruences, identical congruences and residual congruences. Two polynomials $\phi_1(x)$ and $\phi_2(x)$, with integral coefficients, are said to be *residually congruent* with respect to the field $GF[p^n]$ if $\phi_1(x) \equiv \phi_2(x)$ whenever x is any mark of the field; if, however, $\phi_1(x) \equiv \phi_2(x)$ when x is any number whatsoever, then $\phi_1(x)$ is said to be *identically congruent* to $\phi_2(x)$. For example, $x^{p^n} \equiv x$ is a residual congruence, and $(p+1)x \equiv x$ is an identical congruence.

Moreover, if ϕ_1 and ϕ_2 are polynomials in two sets of variables, the x_i and y_i, then we might have a congruence which is residual as far as the x's are concerned, but which is identical as far as the y's are concerned. For example, when C is a modular covariant, (3) in § 2 is a residual congruence as far as the a's are concerned but is an identical congruence as far as the x's are concerned. If, on the other hand, C is a formal covariant, then (3) is an identical congruence with respect to both the a's and the x's. In each case, however, (3)

* *Madison Colloquium Lectures*, p. 42.

is a residual congruence with respect to the ξ's and η's. Since there is no simple notation which will indicate those variables with respect to which a congruence is residual, we shall indicate them in words directly after each congruence.

8. **Manner in which covariants are transformed.** An algebraic covariant, C, is usually defined as a function such that its transform, C', is given by the formula $C' = D^w C$ where D is the determinant of the general transformation of the group; but it is often defined as a function such that $C' = MC$ where M depends merely on the coefficients of the transformation, and the multiplier is then proved to be a power of D. Now it has been customary thus far to define a formal modular covariant in a manner analogous to the first of these methods, but, in this paper, we shall use the second definition and then prove it is equivalent to the first.

Although the definition of a formal modular covariant says nothing explicitly about the nature of its transform except for values of the ξ's and η's in the field, let us see if this indirectly imposes any restriction on the actual form of the transform before we reduce the exponents of the ξ's and η's by Fermat's theorem. For example, under the transformation (1), $L_1 = \alpha_1^{p^n}\alpha_2 - \alpha_1\,\alpha_2^{p^n}$ is replaced by the same function of the primed letters, which is identically equal to

$$D_1\,\alpha_1^{p^n+1} + D_2\,\alpha_1^{p^n}\alpha_2 - D_3\,\alpha_1\,\alpha_2^{p^n} + D_4\,\alpha_2^{p^n+1},$$

where

$$D_1 = \xi_1^{p^n}\eta_1 - \xi_1\,\eta_1^{p^n}, \qquad D_2 = \xi_1^{p^n}\eta_2 - \eta_1^{p^n}\xi_2,$$
$$D_3 = \xi_1\,\eta_2^{p^n} - \xi_2^{p^n}\eta_1, \qquad D_4 = \xi_2^{p^n}\eta_2 - \xi_2\,\eta_2^{p^n}.$$

It will be observed that each D_i is a formal modular invariant of the two pairs (ξ_1, η_1) and (ξ_2, η_2) when they are transformed cogredicntly, and that $D_1 \equiv D_4 \equiv 0$, $D_2 \equiv D_3 \equiv D$, when the ξ's and η's are all in the field.

Let C be any formal covariant of the system S. Then, under (1), C formally —i.e., before any reduction is made modulo p—goes into

$$C(a'; x') = \Sigma D_i(\xi, \eta)\,P_i(a; x) \tag{4}$$

where each D_i* and each P_i is a polynomial in its arguments. If (1) is followed by the transformation

$$x_i' = e_i\,x_1'' + f_i\,x_2'' \qquad (i = 1, 2),$$

where

$$\Delta = e_1 f_2 - e_2 f_1 \not\equiv 0 \quad \text{(in the field)},$$

this is equivalent to applying to the original forms of S the single transformation

$$x_i = \xi_i'\,x_1'' + \eta_i'\,x_2'' \qquad (i = 1, 2), \tag{1'}$$

* These multipliers, D_i, will depend on the covariant, C. When C is L, then the D's are as given in the preceding paragraph.

where the two pairs (ξ_1', η_1') and (ξ_2', η_2') are obtained respectively from the pairs (ξ_1, η_1) and (ξ_2, η_2) by applying the transformation of matrix

$$\begin{pmatrix} e_1 & e_2 \\ f_1 & f_2 \end{pmatrix}. \tag{5}$$

Accordingly, if in (4) we replace the a' and the x' by the corresponding a'' and x'', and apply the transformation (5) to the pairs (ξ, η), then we get an equation of the type (4) but formed for the new transformation (1'). Thus, before any reduction is made modulo p,

$$C(a''; x'') = \Sigma D_i(\xi', \eta') P_i(a; x); \tag{4'}$$

and also

$$C(a''; x'') \equiv MC(a'; x') \quad \text{(in the field)}, \tag{6}$$

whenever the e's and f's are in the field. Both (4) and (4') are ordinary equations, but (6) is a congruence in the field, being an identical congruence with respect to the a's and x's. Combining these, we have

$$D_i(\xi', \eta') \equiv MD_i(\xi, \eta)$$

as an identical congruence in the ξ's and η's. Hence we have proved

THEOREM II. *If C is any formal modular covariant of a system, S, of binary forms with respect to the Galois field $GF[p^n]$, of order p^n, then under any linear transformation (1) of the group, C is formally replaced by an expression of the form $\Sigma D_i(\xi, \eta) P_i(a; x)$ where each P_i is a polynomial in the a's and x's and each D_i is a formal modular invariant of the two cogredient pairs (ξ_1, η_1) and (ξ_2, η_2).*

In a later section, we shall prove that each D_i which is not congruent to zero when the ξ's and η's are in the field is congruent to a power of D (the determinant of the transformation) whenever the ξ's and η's are marks of the field. For isobaric covariants, see Lemma 4 (§ 10); for pseudo-isobaric covariants, see § 16.

Note that this proof holds equally well for classic covariants, and thus, both for classic covariants and modular covariants, the theory of covariants of a system of binary forms has its foundations in the theory of invariants of two cogredient points. The essential differences between the modular and classic theories are, accordingly, associated with two facts: (1) in the classic case, all invariants of two cogredient points are polynomials in one invariant, the determinant $\xi_1\eta_2 - \eta_1\xi_2$; whereas, in the modular case, several other types arise; (2) in the classic case there is no invariant, other than zero itself, which vanishes whenever the ξ's and η's are in the field; whereas, in the modular case, there are invariants of the two pairs which are not identically congruent to zero and yet vanish whenever the ξ's and η's are in the field. From (1)

we see why the transform of a modular covariant is not in general identically a multiple of the original covariant, and from (2) we see why the transform of a modular covariant in general contains terms having no counterpart in the original covariant, before Fermat's theorem is applied. Both remarks are illustrated by L at the beginning of this section.

9. **Remarks on isobarism.** Another fundamental difference between algebraic and modular covariants was mentioned in the introduction and will now be discussed briefly. An algebraic covariant is always isobaric—i.e., all its terms have the same weight.* In fact, an elementary way of determining classic invariants of $f(a; x)$ is to write down a general linear combination of all possible terms having the same degree in the a's and the same weight and then determine what the coefficients of these terms must be in order that the expression be an invariant of f.† But a modular covariant is not in general isobaric, though the weights of any two of its terms can differ at most by a multiple of $p^n - 1$.

These differences are well shown by the algebraic invariants of

$$f = a_0 x_1^2 + 2a_1 x_1 x_2 + a_2 x_2^2$$

and the formal invariants of the same form, modulo 3. All algebraic invariants of f are polynomials in‡

$$\Delta = a_1^2 - a_0 a_2,$$

and the modular invariants are polynomials in Δ, J, B and Γ (see § 6). It will be noticed that of the modular invariants, only one is isobaric and that is the one which is also an algebraic invariant of f; the weights of any terms of the others differ at most by multiples of $p^n - 1 = 2$.

Isobaric covariants

10. **Preliminary lemmas.**

Lemma I. *Let C be a formal modular covariant of a system, S, of binary forms and a number of cogredient points under a group G with coefficients in the Galois field $GF[p^n]$, of order p^n. Then, if C is isobaric, it is invariant modulo p under any transformation* (1) *where the ξ's and η's are independent variables.*

For, since C is invariant modulo p under the transformation

$$x_1 = x_1' + x_2',$$
$$x_2 = \quad x_2',$$

* In the binary form $f = a_0 x_1^n + a_1 x_1^{n-1} x_2 + \cdots + a_n x_2^n$, the weight of the coefficient a_i is defined to be i, and the variables x_1 and x_2 are assigned the weights 1 and 0 respectively. See any standard book on algebraic invariants, such as Elliott, *Algebra of Quantics*, first edition, pp. 36, 48; Dickson, *Algebraic Invariants*, pp. 31, 38.

† For examples, see Elliott, pp. 125–126; Dickson, pp. 36–37.

‡ Dickson, *Algebraic Invariants*, pp. 48, 84; Elliott, loc. cit., pp. 98–99.

and is isobaric, it is formally invariant, modulo p, under any transformation

$$x_1 = x_1' + kx_2', \qquad x_2 = kx_2',$$

where k is a general non-zero scalar. Again using the isobarism, it follows from this last that C is invariant modulo p under the transformation

$$x_1 = x_1' + kx_2', \qquad x_2 = x_2'. \tag{7}$$

Also, C satisfies the condition for covariancy under the transformation

$$x_1 = x_2', \qquad x_2 = x_1', \tag{8}$$

without using Fermat's theorem. Now, using (7) and (8) in conjunction with the isobarism, we prove the lemma.

LEMMA II. *If a formal modular covariant is isobaric, then every one of its factors is necessarily isobaric.*

For, let C_1 and C_2 be any two formal modular covariants which are not isobaric. Then the totality of terms in $C_1\,C_2$ which have the greatest (least) weight is simply the product of the terms of C_1 which have the greatest (least) weight by the terms of C_2 which have the greatest (least) weight. The lemma follows at once.

LEMMA III. *When a formal modular covariant C is isobaric, each of the multipliers $D_i\,(\xi, \eta)$ of Theorem* II *is also isobaric.*

For, by Theorem II, under any transformation (1), we have (3) holding formally as to the a's, x's, ξ's and η's. Since C is isobaric, it is invariant whenever x_1 or x_2 is multiplied by ρ, where ρ is any non-zero scalar. This means that the right member of (3) is homogeneous in ξ_1 and η_1, and also in ξ_2 and η_2. Being also invariant whenever we interchange the x's, the right member of (3) must be of the same degree in ξ_1 and η_1 that it is in ξ_2 and η_2. Hence, if w_1, w_2, w_3 and w_4 are the degrees of any term of some D_i in ξ_1, η_1, ξ_2 and η_2, respectively, then $w_1 + w_2 = w_3 + w_4$. Similarly, since C is invariant formally whenever x_1' or x_2' is multiplied by ρ, we have $w_1 + w_3 = w_2 + w_4$. Therefore $w_1 = w_4$ and $w_2 = w_3$.

LEMMA IV. *When a formal modular covariant, C, is isobaric, each of the multipliers D_i of Theorem* II *is either identically congruent to zero or to a power of the determinant, D, of the transformation.*

By Lemma III, D_i is a linear combination of terms of the form $\xi_1^k\,\eta_1^{w-k}\,\xi_2^{w-k}\,\eta_2^k$, where w is constant for a given covariant, C; and thus each D_i which is not congruent to zero for every transformation of the group is of the form*

* For if D_i is not congruent to zero when the ξ's and η's are in the field, then it must be equal to unity for the identical transformation and hence must contain a term in $\xi_1^w\,\eta_2^w$ with coefficient unity.

$$D_i = (\xi_1 \eta_2 - \eta_1 \xi_2)^w + \Delta_i, \tag{9}$$

where Δ_i is divisible by $\eta_1 \xi_2$ and vanishes whenever the ξ's and η's are in the field. But D_i and $\xi_1 \eta_2 - \eta_1 \xi_2$ are both formal modular invariants of the two pairs, (ξ_i, η_i), and are also isobaric; hence Δ_i is likewise. Being divisible by η_1 and ξ_2, Δ_i is divisible by $L_i = \xi_i^{p^n} \eta_i - \xi_i \eta_i^{p^n}$ $(i = 1, 2)$. By Lemma II, this is impossible unless Δ_i is identically zero in the field. Finally, each D_i which is congruent to zero whenever the ξ's and η's are in the field is of the same type as the Δ_i above. Thus the lemma is proved.

11. **Isobaric seminvariants in terms of the roots.** If a formal modular seminvariant I of a system, S, of binary forms with respect to the field $GF[p^n]$ is isobaric, then, by Lemma I, it is invariant, modulo p, under any transformation

$$\begin{aligned} x_1 &= x_1' + M x_2', \\ x_2 &= x_2', \end{aligned} \tag{10}$$

where M is a general scalar. Hence, as in the theory of algebraic seminvariants, we can prove

LEMMA V. *Let I be any isobaric formal modular seminvariant of a system, S, of binary forms with respect to the Galois field, $GF[p^n]$, of order p^n which is of degree d and weight w. Then I is congruent, modulo p, to the product of a_0^d by a symmetric polynomial, S, in the symbolic ratios α_2/α_1, β_2/β_1, $\cdots$ which is homogeneous in these ratios of degree w. Moreover, S is expressible as a polynomial in the differences of these ratios.*

The proof proceeds as in the theory of classic seminvariants, but (for the sake of completeness) we shall reproduce it for the case when S contains only one form. Since I is isobaric, it can not contain a_0 as a factor, by Lemma II. Hence I is a_0^d multiplied by a polynomial, S, of degree d in a_1/a_0, a_2/a_0, $\cdots$, a_m/a_0. Thus S is a polynomial in the elementary symmetric functions of the symbolic ratios

$$A = \frac{\alpha_2}{\alpha_1}, \frac{\beta_2}{\beta_1}, \frac{\gamma_2}{\gamma_1}, \cdots,$$

and is of degree d in each such ratio; and since I is isobaric, S is homogeneous in these symbolic ratios of total degree w.

But, since I is formally unaltered, modulo p, under transformation (10), it is formally unaltered, modulo p, when the symbolic ratios are diminished by any scalar, M. Accordingly, set

$$\frac{\beta_2}{\beta_1} = \left(\frac{\beta_2}{\beta_1} - \frac{\alpha_2}{\alpha_1}\right) + A,$$

and similarly with the other ratios; and thus

$$S = \left(\frac{\alpha_2}{\alpha_1}\right)^k A_k + \left(\frac{\alpha_2}{\alpha_1}\right)^{k-1} A_{k-1} + \cdots + A_0, \tag{11}$$

where each A_i is a polynomial in the differences such as $(\beta_2/\beta_1) - (\alpha_2/\alpha_1)$, $(\gamma_2/\gamma_1) - (\alpha_2/\alpha_1)$, $\cdots$. As in classic invariant theory, this is impossible unless the A_i $(i \geqq 1)$ are all zero in the field. Hence the lemma is proved. The converse is also readily proved, but will not be needed here.

12. **Isobaric invariants in terms of the ratios.** When I is an isobaric formal modular invariant of a binary form, f, then (by Lemma I) its transform under any transformation (1) when the ξ's and η's are general scalars is connected with I by the relation (2), where (2) is an identical congruence as to the a's and thus is essentially of the same type as the corresponding relation for an algebraic invariant of f. Hence, as in the theory of algebraic invariants, we can prove for invariants the result which is a generalization of Lemma V. A similar argument obtains for an invariant of a system of binary forms and hence for covariants of a system of forms.* Thus we have

LEMMA VI. *Let C (a formal modular covariant of a system S of binary forms, f_i) be of degree d_i in the coefficients of f_i and of degree d in the variables, x_1 and x_2, and let it be isobaric of weight w.*

I. *Then C is identically congruent, modulo p, to $\Pi a_0^{i} x_1^{i}$ multiplied by a linear combination, K, of products of the expressions $(x_1/x_2) + (\alpha_2/\alpha_1)$, and of differences of the type $(\beta_2/\beta_1) - (\alpha_2/\alpha_1)$, where the α's and β's may be symbols arising from the same form, f_i, or may be symbols arising from two different forms.*

II. *Moreover, K is homogeneous in the symbols for each form, f_i, and such that each ratio α_2/α_1 for this form, f_i, occurs in exactly d_i factors in each product.*

III. *Finally, K is symmetric, modulo p, in these symbols for each form f_i.*

In applying the results of §§ 11 and 12, the reader must remember that, although the seminvariant (or invariant) is always symmetric in the pairs (α_1, α_2), (β_1, β_2), $\cdots$ in the form in which it first appears, yet the polynomial, K, in the differences $(\beta_2/\beta_1) - (\alpha_2/\alpha_1)$ is not in general symmetric in the ratios α_2/α_1, β_2/β_1, $\cdots$, but is merely symmetric modulo p. For the binary quadratic $f = a_0 x_1^2 + a_1 x_1 x_2 + a_2 x_2^2 = \alpha_x \beta_x$ has the formal modular invariant $a_1 = \alpha_1 \beta_2 + \alpha_2 \beta_1 \equiv \alpha_1 \beta_1 [(\beta_2/\beta_1) - (\alpha_2/\alpha_1)]$, modulo 2. The first is actually symmetric; but the second is not symmetric, though it is symmetric modulo $p = 2$.

13. **Isobaric formal covariants as algebraic covariants.** From Lemma VI it is evident that, if we express a_0 for each form, f_i, in terms of the symbols, we have

THEOREM III. *Every isobaric formal modular covariant, C, of a system of binary forms is expressible as a polynomial in symbolic covariants of the type $\alpha_x = \alpha_1 x_1 + \alpha_2 x_2$ and in determinantal symbolic invariants of the type*

* For every formal covariant of the system S is a formal modular invariant of the system S' consisting of the forms of S and the additional linear form $l = X_2 x_1 - X_1 x_2$ in which X_1, X_2 have been replaced by x_1, x_2, respectively. The proof of this is the same as for algebraic covariants.

$$(\alpha\beta) = \alpha_1\beta_2 - \alpha_2\beta_1.$$

But it is well known that the product, P, of $x_2^q \Pi a_0^{q_i}$ by any expression, K, in the symbols satisfying the conditions I and II of Lemma VI is an algebraic covariant of the system S of forms, though not necessarily a rational one.* A necessary and sufficient condition that P be rational in the coefficients of f_i is that K be actually symmetric in the ratios β_2/β_1, α_2/α_1, $\cdots$ formed for f_i. This is stronger than condition III of Lemma VI, as is illustrated by the symbolic invariant $(\alpha\beta)$ for the binary quadratic, mod 2 (see § 12). The kth power of $(\alpha\beta)$ is symmetric, mod 2, for any k; but is actually symmetric only when k is even. Thus we have

COROLLARY 1. *Every isobaric formal modular covariant of a system, S, of binary forms with respect to the Galois field, $GF[p^n]$ of order p^n, is congruent, modulo p, to an algebraic covariant of S, either rational or irrational.*

But every irrational algebraic covariant of S is known to be a root of an algebraic equation whose coefficients are rational, algebraic covariants of S, and thus we have

COROLLARY 2. *Every isobaric formal modular covariant of S is congruent, modulo p, to a root of an algebraic equation whose coefficients are rational algebraic covariants of S.*

PSEUDO-ISOBARIC COVARIANTS

14. **Informal discussion.** In the preceding sections, we have considered formal modular covariants that are actually isobaric and have seen that, in many ways, such a covariant is much the same as an ordinary classic covariant. But if a formal modular covariant is not isobaric, what then? For example, consider the formal modular invariants of a point with respect to the field $GF[p^n]$. Any such invariant, I, whose degree, d, is less than $p^n + 1$ must be an ordinary invariant of the point. For, since the weights of any two terms would have to be congruent, mod $p^n - 1$, I would have to be of the form $Ax_1^{p^n-1} + Bx_2^{p^n-1}$ where A and B are constants. But, if the degree is $< p^n$, there is no opportunity to apply Fermat's theorem to the coefficients of the transformation, and thus I must be invariant in the classic sense and therefore identically zero in the field. If $p^n < d \equiv p^{2n}$, it can be shown by a

* Dickson, *Algebraic Invariants*, p. 55, ex. 7; Elliott, loc. cit., p. 93. (Some people do not agree with my statement, and (in justification) say that P in general has no actual meaning in the a's. But any polynomial, P, in the symbols which is not actually symmetric in the symbols, $\alpha_1, \alpha_2, \beta_1, \beta_2, \cdots$ for a particular form, f_i, is a root of an algebraic equation whose coefficients are the elementary symmetric functions of P and the other symbolic invariants obtained from P by interchanging any two pairs of symbols for this form f_i; and hence P is a root of an algebraic equation whose coefficients are polynomials in the coefficients of f_i. Similarly, P is a root of an algebraic equation whose coefficients are polynomials in the coefficients of S, and thus P is an algebraic function of the coefficients of S, though it is not necessarily an explicit algebraic function of these coefficients. Hence why is it not an algebraic covariant of S?)

similar argument that either I is an ordinary invariant of the two cogredient pairs, (x_1, x_2), $(x_1^{p^n}, x_2^{p^n})$, and therefore $I = kL$ where k is a constant and $L = x_1^{p^n} x_2 - x_1 x_2^{p^n}$; or LI is an ordinary invariant of the two cogredient pairs (x_1, x_2) and $(x_1^{p^{2n}}, x_2^{p^{2n}})$, and therefore $LI = k(x_1^{p^{2n}} x_2 - x_1 x_2^{p^{2n}})$. In fact, it is well known that all formal modular invariants of the point (x_1, x_2) are polynomials in L and $Q = (x_1^{p^{2n}} x_2 - x_1 x_2^{p^{2n}})/L$. This might be stated more strikingly by saying that every such invariant is a rational, integral algebraic invariant of three points, or is a quotient of two such invariants. Moreover, a similar statement is true of the formal modular invariants of a pair of cogredient points, (x_1, x_2) and (y_1, y_2).* Is this true more generally?

15. **Preliminary lemmas.** Consider a formal modular invariant, $I = \sum I_j$, of the binary form $f(a; x)$ which is of degree d and which is not isobaric. In particular, let I_0 be a term of I whose weight, w_0, is less than or equal to that of any other term of I, and let I_1 be any other term of I. Then, since I is pseudo-isobaric, the weight of I_1 will be $w_1 = w_0 + l(p^n - 1)$ where l is an integer, positive or zero. When the coefficients of $f(a; x)$ are replaced by the corresponding functions of the symbols α_1, α_2, β_1, β_2, $\cdots$, I is homogeneous of degree md in these symbols. Also, if I_j is of weight w_j, it is of degree $md - w_j$ in symbols with subscript 1 and of degree w_j in symbols with subscript 2. Their difference, $md - 2w_j$, we shall call the excess, as in classic invariant theory.

LEMMA VII. *The excess of any term of a formal modular covariant is congruent to zero, modulo* $p^n - 1$.†

For, by interchanging the subscripts 1 and 2 throughout I, it follows that I contains a term whose degree in the symbols with subscript 1 is w_j and whose degree in the symbols with subscript 2 is $md - w_j$, and therefore of weight $md - w_j$. Hence $md - w_j \equiv w_j \pmod{p^n - 1}$, and the excess of any term of I is a multiple of $p^n - 1$, say $E_j(p^n - 1)$.

Now I^{p^n} is identically congruent in the field to $\sum I_j^{p^n}$,‡ and thus the latter is a formal modular invariant of $f(a; x)$. It may also be regarded as an invariant of $f(a^{p^n}; x)$, but we shall show that it is better for our purpose to regard it as a simultaneous invariant of $f(a; x)$ and $f(a^{p^n}; x)$.

For in I^{p^n} the term $I_0^{p^n}$ is of excess $E_0 p^n (p^n - 1)$, and hence, if we replace $E_0(p^n - 1)$ factors of the form $\alpha_1^{p^n}$ by as many factors of the form $\bar{\alpha}_1$, the excess of the resulting symbolic expression is $E_0(p^n - 1)$. If $w_0 \geqq E_0$, then by replacing E_0 additional factors of the form $\alpha_1^{p^n}$ by as many factors α_1,

* W. C. Krathwohl, *Modular invariants of two pairs of cogredient variables*, American Journal of Mathematics, vol. 36 (1914), pp. 449–460.

† After finishing this MS., I discovered that Professor Glenn stated and proved this in *Concerning an analogy between formal modular invariants and the class of algebraic invariants called Booleans*, American Journal of Mathematics, vol. 37 (1915), p. 75.

‡ Dickson, *Linear Groups*, p. 15.

the excess of the resulting symbolic expression is zero. $I_0^{p^n}$ is now of degree $E_0 p^n$ in $\overline{\alpha}$'s having subscript 1 and of degree zero in $\overline{\alpha}$'s having subscript 2; also, it is of degree $(w_0 - E_0) p^n$ in α's having subscript 1 and of degree $w_0 p^n$ in the α's having subscript 2. Here and elsewhere in this work, we use the expression "symbols α's" to denote any and all symbols such as $\alpha_1, \alpha_2, \beta_1, \gamma_1, \cdots$; similarly with the expression "symbols $\overline{\alpha}$."

Moreover, if $w_0 \geqq E_0$, the same device will serve to make all terms of I^{p^n} of excess zero and also to make all terms of the same weight, w_0. For the weight of any term, I_ν, of I is $w_0 + k(p^n - 1)$ and hence is of excess $(E_0 - 2k)(p^n - 1)$ which we shall assume positive for convenience. Accordingly, if in $I_\nu^{p^n}$ we replace $(E_0 - 2k)(p^n)$ factors of the form $\alpha_1^{p^n}$ by as many factors $\overline{\alpha}_1$, the result is of excess zero, but of weight $[w_0 + k(p^n - 1)] p^n$. Next replace kp^n factors $\alpha_1^{p^n}$ by as many factors $\overline{\alpha}_1$ and replace kp^n factors $\alpha_2^{p^n}$ by as many factors $\overline{\alpha}_2$, and we get an expression whose excess is zero and weight is $w_0 p^n$. This is legal, since $w_0 \geqq E_0 - k + 1$ and $w_0 \geqq k$. In case the excess of I_ν is negative, interchange the subscripts 1 and 2 in the foregoing by an argument similar to that used in Lemma VII. Thus the remark at the beginning of this paragraph is proved.

For convenience, we shall speak of the α's, β's, $\cdots$ (the symbolic coefficients of the linear factors $\alpha_x, \beta_x, \cdots$ of $f(a; x)$) as the *primary symbols*; and we shall speak of the totality of the original symbols $\alpha_1, \alpha_2, \beta_1, \beta_2, \cdots$ together with the new symbols $\overline{\alpha}_1, \overline{\alpha}_2, \overline{\beta}_1, \overline{\beta}_2, \cdots$, and any other symbolic coefficients introduced in the above manner as the *secondary symbols*.

We see, accordingly, that every pseudo-isobaric formal modular invariant, I, of $f(a; x)$ is such that a suitable power of I is isobaric in the coefficients of the symbolic linear factors of $f(a; x)$ and of $f(a^{p^n}; x)$, provided $0 \neq w_0 \geqq E_0$. Even if $0 \neq w_0 < E_0$, this argument shows that the excess of I^{p^n} in the secondary symbols is $(E_0 - w_0)(p^n - 1)$ less than that of the original I. But, since the excess is finite, it follows that, by repeating this process at most $[E_0/w_0]$* times, we prove Lemma VIII for the case $w_0 \neq 0$. It will be observed, however, that for this purpose we have to introduce another set of new symbols, say $\overline{\alpha}_1, \overline{\alpha}_2, \overline{\beta}_1, \overline{\beta}_2, \cdots$, and possibly still other sets having three or more dashes. These, together with those introduced above, we group under the heading "secondary symbols."

When $w_0 = 0$, we first multiply I by any invariant which is not an absolute invariant and then proceed as above. Thus we have

LEMMA VIII. *Let I be a pseudo-isobaric invariant of a binary form*

$$f(a; x) = \Pi \alpha_x$$

with respect to the Galois field, $GF[p^n]$, of order p^n; and let the lowest weight of all

* Here $[E_0/w_0]$ means the smallest integer which is not less than E_0/w_0.

the terms of I be denoted by w_0. Then, if $w_0 \neq 0$, some finite power of I is congruent to a polynomial in the secondary symbols for f which is isobaric. If $w_0 = 0$, then some power of I is congruent to the quotient of two polynomials in the secondary symbols each of which is isobaric For examples, see Part IV.

16. **Pseudo-isobaric modular covariants as algebraic covariants.** Let J denote that formal modular invariant of f which is obtained as a result of applying the method of Lemma VIII to the pseudo-isobaric invariant I and which is, accordingly, isobaric of weight w in the secondary symbols. If we apply to J the transformation (1), then the transform $J' = J(\alpha')$ is identically equal to

$$J(\alpha'; \bar{\alpha}'; \bar{\alpha}'; \cdots) = \sum D_k(\xi, \eta) P_k(\alpha; \bar{\alpha}; \bar{\alpha}; \cdots) + \tau \tag{12}$$

where each D_k (i) is a formal modular invariant of the two pairs (ξ_1, η_1) and (ξ_2, η_2), (ii) is congruent to zero or to D^w whenever the ξ's and η's are all in the field, (iii) is isobaric of weight w in the ξ's and η's; and where τ is a polynomial in the ξ's and η's and in the secondary symbols which vanishes whenever the dashed symbols are replaced by the proper powers of the corresponding primary symbols. That is, J is not in general an invariant function of the secondary symbols when these symbols are taken as independent, but it is an invariant function when the dashed symbols are replaced by the proper expressions in the primary symbols.

Since J is isobaric in the secondary symbols, Lemmas I–IV apply to J. If we proceed to Lemma V, we see that the proof applies to J with the exception that, since J is not in general rational in the coefficients of $f(a; x)$, $f(a^{p^n}; x)$, $f(a^{p^{2n}}; x)$, $\cdots$, then J is not in general congruent (modulo p) to the product of a_0^w by a symmetric polynomial, S, in the secondary symbolic ratios; but, nevertheless, J is congruent (modulo p) to a product of secondary symbols with subscript 1 by a polynomial, S, in the secondary symbolic ratios. Hence, as before, S is expressible as a polynomial in the differences of the secondary symbolic ratios. Thus the proof of Lemma VI applies to J with the restrictions noted.

Hence we prove Theorem III and its corollaries for J, and we have

THEOREM IV. *If I is a formal modular covariant of a system of binary forms with respect to the Galois field $GF[p^n]$ of order p^n, then I^q (where q is a sufficiently large integer) is congruent to one of the two following forms*:

(1) *a polynomial in symbolic covariants of the types α_x, α_X, A_x, $(\alpha\beta)$, $(A\beta)$, etc., where $X = x^{p^n}$ and $A = \alpha^{p^n}$*;

(2) *a quotient of two polynomials described in* (1).

Since these symbolic expressions are (symbolic) algebraic covariants of $f(a; x)$, $f(a^{p^n}; x)$, etc., we have

COROLLARY 1. *Every pseudo-isobaric formal modular covariant, C, of a system S of binary forms with respect to the Galois field $GF[p^n]$ is congruent,*

modulo p, to a symbolic algebraic covariant, C, of the forms of S and certain related forms obtained from the forms of S by replacing each coefficient of these forms by its p^nth power.

In particular, this means that the function J may be chosen in such a way that, under the transformation (1), J goes into J', where J' is equal to (not merely congruent to) $D^w J$, when J is expressed as an isobaric function of the secondary symbols. Note, however, that J is not necessarily rational in the coefficients of the forms of S and the related forms. These remarks are illustrated by examples given in Part IV.

17. **Finiteness of determinantal symbols.** In the preceding section, we proved that all formal modular invariants I of a binary system, S, are congruent, modulo p, to rational functions in symbolic invariants of S of a certain special type called determinantal symbols. That is, I is congruent, modulo p, to a rational function of symbolic invariants of the types $(\alpha\beta)$, $(\alpha^{p^r}\alpha)$ and symbols obtained from these by replacing a pair of symbols by their p^nth powers one or more times. Conceivably, as the degree of I increases, the degree of the requisite determinantal symbols might increase without bound, so that we would need an infinite number of determinantal symbols to obtain all formal modular invariants of S.

But, in Part II, we proved that all formal modular invariants of S are congruent to polynomials in a finite number of symbolic invariants which are formal modular invariants of the (symbolic) coefficients of the symbolic linear factors of the forms of S. Now apply Theorem IV to these symbolic invariants, and we see that all symbolic invariants of S are congruent to rational functions of determinantal symbols. Hence we have

THEOREM V. *All formal modular covariants of a system of binary forms are congruent, in the field, to rational functions of a finite number of symbolic covariants of the types described in Theorem* IV.

In fact, this theorem could have been proved directly in much the same manner as Theorem IV was proved and then Theorem IV would have followed as a corollary of Theorems I and V.

PART IV. EXPRESSION OF COVARIANTS IN SYMBOLIC FORM

18. **Method for algebraic covariants.** In the symbolic theory of algebraic covariants, after proving that every such covariant is expressible as a polynomial in symbols of the two types, $(\alpha\beta)$ and α_x, there is then given a clear-cut method for expressing any covariant in terms of these symbolic covariants. For completeness, this method will be reproduced here, as we shall need to refer to several of the results.

Let I be any algebraic invariant, of weight w, of a system of forms whose coefficients are the a's; and let $I(a')$ be its transform under (1). Under this

transformation, α_1 and α_2 go into α_ξ and α_η, respectively; and similarly for the β's, γ's, etc. Just as the a's are polynomials in $\alpha_1, \alpha_2, \beta_1, \beta_2, \cdots$, so their transforms are polynomials in $\alpha_\xi, \alpha_\eta, \beta_\xi, \beta_\eta, \cdots$. Hence, in the equation (2),

(i) replace every a' by the corresponding polynomial in α_ξ, α_η, etc.;

(ii) operate on both sides of the resulting equation w times with

$$V = \frac{\partial^2}{\partial\xi_1\,\partial\eta_2} - \frac{\partial^2}{\partial\xi_2\,\partial\eta_1};$$

(iii) divide both sides of the result of (ii) by $(w+1)!\,w!$.

From the left we get a polynomial in the symbolic invariants of the type $(\alpha\beta)$, and from the right we get $I(a)$.

In proving this, use is made of the following:

LEMMA A.

$$V^w(\xi\eta)^w = (w+1)!\,w!.$$

LEMMA B. *The result of operating by V^w on a product of k terms of type α_ξ and l terms of the type β_η is a sum of terms each containing $k-w$ factors α_ξ, $l-w$ factors β_η and w factors $(\alpha\beta)$.*

19. **Difficulties in modular case.** It is at once evident that this method is not applicable directly to formal modular invariants, since the divisor in (iii) is congruent to zero in the field whenever $w \geqq p-1$. This difficulty, of course, might be superficial in the sense that it might be due to some accidental peculiarity in the proof and not to any essential characteristic of the problem; but, as will be apparent presently, there is a deep-seated difficulty.

Lemma B is proved by direct verification combined with induction, and thus is lost the full significance of the reason why it is true. Now each symbolic expression α_ξ and each β_η is an invariant of the two points (ξ_1, ξ_2) and (η_1, η_2), which are transformed cogrediently when the original variables are subjected to any linear transformation

$$\begin{aligned} x_1 &= \kappa\bar{x}_1 + \lambda\bar{x}_2, \\ x_2 &= \mu\bar{x}_1 + \nu\bar{x}_2, \end{aligned} \qquad \Delta = \kappa\nu - \lambda\mu \neq 0, \tag{13}$$

thus replacing (1) by

$$\begin{aligned} \bar{x}_1 &= \bar{\xi}_1 x'_1 + \bar{\eta}_1 x'_2, \\ \bar{x}_2 &= \bar{\xi}_2 x'_1 + \bar{\eta}_2 x'_2, \end{aligned} \tag{14}$$

where the $\bar{\xi}$'s and $\bar{\eta}$'s are cogredient with the x's. Moreover, V is an invariantive operator whenever the ξ's and η's are transformed cogrediently. Lemma B is an inevitable consequence of these facts.

If we look at Theorem II (§ 8), we see that in the modular case we have the identity (4) where each D_i is a formal modular invariant of (ξ_1, η_1) and

(ξ_2, η_2) when they are transformed cogrediently; but D_i is not in general a formal modular invariant of the two pairs (ξ_1, ξ_2) and (η_1, η_2), though it is clearly a modular invariant of these pairs. Moreover, the only way to operate on a product of a number of symbols of the type α_ξ and β_η in such a way as (i) to eliminate the ξ's and η's and (ii) to obtain as result a formal modular invariant of the α's, β's, etc., is to use an operator which has the invariantive property when (ξ_1, ξ_2) and (η_1, η_2) are transformed cogrediently. Now the natural analogues of V for the formal modular case are, by Theorem II, those formed on

$$\xi_1^{p^n}\eta_2 - \eta_1^{p^n}\xi_2, \quad \text{etc.},$$

as models—i.e., by replacing each ξ by the operator indicating partial differentiation with respect to that ξ, and similarly with the η's—but they can not give the desired results since these operators are not invariantive when the ξ's are transformed cogrediently with the η's.

20. **Significance of results of Part III.** Such is the essential difficulty of the problem and such is one of the reasons for proving the theorems of Part III. There it is shown that (i) if C is isobaric, then each D_i is an algebraic invariant of the two pairs (ξ_1, η_1) and (ξ_2, η_2) and hence is an invariant of the two pairs (ξ_1, ξ_2) and (η_1, η_2); (ii) if C is not isobaric, then a suitably high power of C, say C^q, is such that, when (2) is formed for C^q, then each D_i is an algebraic invariant of the two pairs (ξ_1, η_1) and (ξ_2, η_2) and hence is an invariant of the pairs (ξ_1, ξ_2) and (η_1, η_2). In view of the remarks of § 19, these results smooth out some of the difficulties.

21. **First method.** Let I be any formal modular invariant of a system S of binary forms with respect to the Galois field, $GF[p^n]$, of order p^n, and let J be a formal modular invariant of S which is obtained from I by applying the method of Lemma VIII and which is isobaric in the secondary symbols. The first method of expressing J in symbolic form is suggested by the proof of Lemma V (§ 11). Although it lacks originality, it has the advantage that it leads at once to a result without exception or qualification. If J is of degree d_i in the coefficients of f_i, divide J by $\pi\alpha_1^{d_i}$ where the α_1 range over all symbols for f_i having subscript 1. Do this for every form f_i. The quotient, S, is a polynomial in the ratios $\alpha_2/\alpha_1, \beta_2/\beta_1, \cdots, \overline{\alpha_2}/\overline{\alpha_1}, \cdots$ with coefficients which are marks of the field. Single out any particular ratio of the primary symbols, say α_2/α_1, and select all terms of S which do not contain α_2/α_1 as a factor. In the resulting expression, replace β_2/β_1 by $(\beta_2/\beta_1) - (\alpha_2/\alpha_1)$, $\alpha_2^{p^n}/\alpha_1^{p^n}$ by $(\alpha_2^{p^n}/\alpha_1^{p^n}) - (\alpha_2/\alpha_1)$, and similarly for all symbolic ratios. The resulting expression, E, is clearly an invariant, since it is a polynomial in the determinantal symbols. Moreover, it is identically congruent, modulo p, to J, for it is simply A_0 in equation (11).

22. **Second method.** It is usually easy to determine by inspection a product, π, of determinantal symbols in such a way that the terms of π which have the greatest excess are among the terms of J which have the greatest excess; and hence, by a finite number of steps, we can express J in symbolic form. No attempt has been made to crystallize this into a definitive, systematic process, though possibly this might be done.

For example, consider $a_0 x_1^2 + 2a_1 x_1 x_2 + a_2 x_2^2$. Modulo 3, this has the formal modular invariant

$$\begin{aligned} B &= a_1 (a_1 - a_0)(a_1 + a_0)(2a_0 + a_2)(2a_1 + a_2)(a_1 + a_2) \\ &\equiv (\alpha_1\beta_2 + \alpha_2\beta_1)(\alpha_1\beta_2 + \alpha_2\beta_1 + \alpha_1\beta_1)(\alpha_1\beta_2 + \alpha_2\beta_1 - \alpha_1\beta_1) \\ &\qquad \times (-\alpha_1\beta_1 + \alpha_2\beta_2)(\alpha_1\beta_2 + \alpha_2\beta_1 + \alpha_2\beta_2)(\alpha_1\beta_2 + \alpha_2\beta_1 - \alpha_2\beta_2) \\ &\equiv \textstyle\sum \alpha_1^6\beta_1^3\beta_2^3 - \sum \alpha_1^6\beta_1\beta_2^5 + \sum \alpha_1^4\beta_1^3\alpha_2^2\beta_2^3 \\ &\qquad - \textstyle\sum \alpha_1^3\beta_1^2\alpha_2^3\beta_2^4 + \sum \alpha_1^5\alpha_2\beta_2^6 - \sum \alpha_1^3\alpha_2^3\beta_2^6 \\ &\equiv \textstyle\sum (\alpha^3\beta)^2(\beta^3\beta) \qquad (\text{mod } 3). \end{aligned}$$

The terms of greatest excess in B are $\sum \alpha_1^6\beta_1^3\beta_2^3$ which suggests that we use to "kill off" these terms $\sum (\alpha^3\beta)^2(\beta^3\beta)$, since the only way to express $\alpha_1^6\beta_1^3\beta_2^3$ in the secondary symbols $\alpha_1, \alpha_2, \beta_1, \beta_2, \alpha_1^3, \alpha_2^3, \beta_1^3, \beta_2^3$ in such a way that it shall be isobaric in these secondary symbols is $\alpha_1^3\beta_2\,\alpha_1^3\beta_2\,\beta_1^3\beta_2$. Thus we take as π, $(\alpha^3\beta)^2(\beta^3\beta)$. It is then readily discovered that $\sum (\alpha^3\beta)^2(\beta^3\beta)$ accounts for all terms, modulo 3.

23. **Third method.** The preceding method is closely related to the one about to be explained. Although the second method has none of that generality which attracts the artist in pure mathematics, yet it has the advantage in any concrete case, while the present method is to be preferred in any general argument.

Under transformation (1), J is carried into

$$J(a') \equiv D^w J(a) \qquad (\text{mod } p) \tag{15}$$

where this is an identical congruence in the a's and the coefficients of the transformation. Moreover, it is understood that $J(a')$ is so written that it is isobaric of weight w in the secondary symbols as in Lemma VIII. Thus, if each a' is replaced by its expression in terms of the α_ξ and α_η, $\overline{\alpha}_\xi$ and $\overline{\alpha}_\eta$, etc., the left member of (15) is the sum of a number of terms each of which is a product of w linear functions of the ξ's and w linear functions of the η's.

Now operate on both sides of (15) with V^w as given in (ii), § 18. Since the left member is isobaric in the secondary symbols, we can apply Lemma B. Hence, from the left member we get a sum of a number of terms, each of which is a product of w determinantal symbols such as $(\alpha\beta)$, $(\alpha^{p^n}\beta)$, $(\alpha^{p^n}\alpha)$, etc. From the right we get $(w+1)!\,w!\,J(a)$. Call the result equation E.

If $(w+1)!\,w! \equiv 0 \pmod{p}$, then the totality of those terms on the right of equation E whose coefficients are not congruent to zero, modulo p, constitute a syzygy, modulo p. Moreover, since the right member of (15) is symmetric, modulo p, in the primary symbols, this syzygy must be symmetric, modulo p, in these symbols and hence congruent to a function which is rational in the coefficients of the ground forms, f. Hence, if we know the necessary syzygies or use discretion, we can replace equation E by one (i) which is equivalent to E and (ii) which has each coefficient congruent to zero, modulo p. If we now cancel p from both sides of this equation, we have an equation, E_1, of the same type as E, but such that the multiplier on the right is not divisible by so high a power of p. Now treat equation E_1 as we treated E above. These steps are all justified, since J is an algebraic invariant of certain forms.

24. **Illustrations.** A fundamental system of formal modular invariants, modulo 3, of the binary quadratic was expressed in symbolic form in § 6.

It is known that a fundamental set of formal modular invariants, modulo 2, of the cubic

$$a_0 x_1^3 + a_1 x_1^2 x_2 + a_2 x_1 x_2^2 + a_3 x_2^3 = \pi\alpha_x = (\alpha_1 x_1 + \alpha_2 x_2)(\beta_1 x_1 + \beta_2 x_2)(\gamma_1 x_1 + \gamma_2 x_2)$$

is

$$K = a_1 + a_2, \qquad \Delta = a_0 a_3 + a_1 a_2, \qquad I = a_0^2 + a_0 K + \delta_{00},$$

$$k = a_0 \delta_{00}, \qquad g = \beta_2^2 + \beta(\Delta + K^2) + (\Delta + \delta_{00})(\beta + a_0 K + K^2)$$

where

$$\beta = a_0 a_1 + a_1^2,$$
$$\delta_{00} = (a_0 + a_1 + a_2 + a_3) a_3.$$

After a little computation, it is readily seen that

$$\Delta \equiv (\alpha\beta)(\beta\gamma)(\gamma\alpha),$$
$$I \equiv \frac{(\alpha^1\alpha)(\beta^1\beta)(\gamma^1\gamma)}{(\alpha^2\alpha)(\beta^2\beta)(\gamma^2\gamma)} + K^2 + \Delta,$$
$$k = (\alpha^2\alpha)(\beta^2\beta)(\gamma^2\gamma).$$

But every attempt to express K as a polynomial in the determinantal symbols is bound to fail, since it is not possible to express K as an isobaric function of the secondary symbols. In fact, K is the invariant which forced the author to realize Theorem IV and the lemmas that lead thereto. It is easy to show that

$$K^2 \equiv \sum(\alpha^2\gamma)(\beta^2\gamma) + (\alpha\beta)(\beta\gamma)(\gamma\alpha).$$

Similarly, it is not possible to express g as a polynomial in determinantal symbols, though (by Theorem IV) a power of g is so expressible.

The author has also expressed some covariants of these two forms in symbolic form, but the amount of numerical work necessary to do this seemed to increase rapidly with the degree of the covariant called J in Part III. In short, the author would not recommend that the theory of Part III be applied in detail to any but the simplest concrete case, unless there should prove to be some powerful systematic procedure not yet discovered.

Part V. Finiteness theorems

25. **Preliminary.** In Part III, we saw that every formal modular covariant of a system, S, of binary forms with respect to the Galois field $GF[p^n]$, of order p^n, is expressible as a polynomial in a finite number of symbolic covariants. Moreover, if the covariant is rational and integral, it must be such that when expressed in terms of these symbolic covariants it is unaltered, modulo p, when any two pairs of primary symbols, (α_1, α_2) and (β_1, β_2), are interchanged. It must be carefully noted, however, that this is not equivalent to saying that it is symmetric in the symbolic covariants.

Using this result, we can prove that the set of all formal modular covariants of S which are rational and integral possesses the finiteness property—i.e., they are all expressible as polynomials in a finite number of these covariants. The proof given below is Hilbert's proof of the finiteness theorem for algebraic covariants, with only one slight modification. The theorem is proved for invariants of a general binary system, S, since the covariants of a system S are coextensive with the invariants of another system S'.

It is readily seen that Hilbert's beautiful proof,* when shorn of non-essentials, depends entirely on the following properties of algebraic invariants:

I. All invariants are polynomials in a finite number of polynomials, P, in a certain finite set of pairs of auxiliary variables. Since these polynomials, P, are invariants under the group, we shall call them elemental invariants. In the case of algebraic invariants, these auxiliary variables are the coefficients of the symbolic linear factors of the forms of S and the elemental invariants are the determinantal symbolic invariants of type $(\alpha\beta)$.

II. If the invariant is rational and integral, it must be such that, when expressed in terms of the elemental invariants, it is unaltered (with respect to the field of definition) when any two pairs of the auxiliary variables are interchanged.

His proof uses the following theorem about diophantine equations: A system of any number of linear diophantine equations has only a finite number of simple sets of solutions, all other solutions being sums of positive multiples of this finite number of sets of solutions.

* *Ueber die Endlichkeit des Invariantensystems für binäre Grundformen*, Mathematische Annalen, vol. 33 (1889), pp. 223–226.

Now Part II proves that these two properties hold for formal modular invariants where the auxiliary variables are the primary symbols $\alpha_1, \alpha_2, \beta_1, \beta_2, \cdots$ and where the elemental invariants are the finite number of symbolic invariants of Theorem I. Moreover, Hilbert's proof makes no use of properties peculiar to a field of characteristic zero except in one place which is readily avoided. Thus his proof applies here.

26. **Finiteness of covariants.** For convenience, let $I_1, I_2, \cdots, I_\mu$ be a fundamental set of symbolic formal modular invariants of S. Then every formal modular invariant of S is expressible in the form

$$I = \sum M_i I_1^{s_1} I_2^{s_2} \cdots I_\mu^{s_\mu} \tag{16}$$

where the exponents, s_j, are positive integers or zero and the coefficients, M_i, are marks of the Galois field, $GF[p^n]$. Necessary and sufficient conditions that I be a rational and integral function of the coefficients of the forms of S are that (i) when the right member of (16) is expanded, it is a function of the pairs $(\alpha_1, \alpha_2), (\beta_1, \beta_2), \cdots$ which is symmetric, modulo p, (ii) each term of this expansion contains the same number of α's as β's, the same number of β's as γ's, etc.

If $d_j(\alpha)$ denotes the degree of I_j in the pair (α_1, α_2), then the last condition gives a set of linear diophantine equations of which two are

$$\begin{aligned} d &= s_1 d_1(\alpha) + s_2 d_2(\alpha) + \cdots + s_\mu d_\mu(\alpha) \\ &= s_1 d_1(\beta) + s_2 d_2(\beta) + \cdots + s_\mu d_\mu(\beta), \end{aligned} \tag{17}$$

there being one such equation for every pair of symbols. By the theory of linear diophantine equations, it follows that every rational integral invariant, I, is expressible in the form

$$I = \sum N_i C_1'^{t_1} C_2'^{t_2} \cdots C_r'^{t_r}, \tag{18}$$

where each C_i is a symbolic invariant of S which has the property (ii) above and where (iii) there is one C_i corresponding to each simple set of solutions of the set of diophantine equations (17), and (iv) the C_i are the same for all invariants, I, of the system S.

Moreover, if σ_i is the number of the symbolic invariants which are not identically congruent, modulo p, and which are obtained from a particular C_i by interchanging any two pairs of symbols, then C_i satisfies an algebraic equation of degree σ_i where the coefficients of the various terms are (aside from sign) the elementary symmetric functions of C_i and its conjugates. Since C_i and therefore each of its conjugates has the property (ii), these coefficients, K_j, are formal modular invariants of S which are rational and integral in the a's. Hence I may be written

$$I = \sum L_j \left(\sum Q_i C_1'^{t_1} C_2'^{t_2} \cdots C_r'^{t_r} \right) \qquad (0 \leqq t_i < \sigma_i) \tag{19}$$

where the Q's are in the field and where the summation in the parentheses is symmetric, modulo p, in the pairs of symbols (α_1, α_2), (β_1, β_2), etc. Moreover, the L_j are polynomials in a finite set of rational, integral formal modular invariants, viz., the K's.

Hence by referring to the sufficient conditions given above that a symbolic invariant be rational and integral, we have

THEOREM VI. *Let S be a system of binary forms with coefficients which are independent variables and let G be the total group of linear transformations on the variables whose coefficients are marks of the Galois field, $GF[p^n]$, of order p^n. Then all formal modular covariants of S under the group G are expressible as polynomials in a finite number of such covariants.*

27. **Finiteness of syzygies.** By using, as in classic invariant theory, Hilbert's useful theorem about an infinite sequence of polynomials, we can show that every syzygy, $S \equiv 0$, among the covariants of the system S is of the form

$$S \equiv \sum H_j S_j,$$

where (i) the H_j are formal modular covariants of S;

(ii) the S_j are polynomials in the formal modular covariants such that $S_j \equiv 0$ is a syzygy among the formal modular covariants of S;

(iii) the S_j are finite in number and are the same for all syzygies.

Thus we have

THEOREM VII. *The syzygies among the formal modular covariants of the system S possess the finiteness property in the sense that all of them are consequences of a finite number of them, $S_1 \equiv 0$, $S_2 \equiv 0$, $\cdots$, $S_k \equiv 0$.*

MOUNT HOLYOKE COLLEGE,
SOUTH HADLEY, MASS.

ON THE MEAN-VALUE THEOREM CORRESPONDING TO A GIVEN LINEAR HOMOGENEOUS DIFFERENTIAL EQUATION*

BY

G. PÓLYA

I consider the linear homogeneous differential expression of order n

$$Lf(x) \equiv f^{(n)}(x) + \phi_1(x) f^{(n-1)}(x) + \phi_2(x) f^{(n-2)}(x) + \cdots + \phi_n(x) f(x).$$

I suppose $\phi_1(x)$, $\phi_2(x)$, $\cdots$, $\phi_n(x)$ are continuous, $f(x)$ is differentiable n times, the variable x and all the functions in question are real. I then prove some theorems concerning the equation

$$Ly \equiv \frac{d^n y}{dx^n} + \phi_1 \frac{d^{n-1} y}{dx^{n-1}} + \cdots + \phi_n y = 0 \tag{1}$$

which are, I believe, new, but there is one special case of them which belongs to the most classical part of the differential calculus. This is the case of the simplest equation of the form (1), that is

$$\frac{d^n y}{dx^n} = 0. \tag{2}$$

Denote the Wronskian

$$\begin{vmatrix} f_1(x) & f_1'(x) & \cdots & f_1^{(n-1)}(x) \\ f_2(x) & f_2'(x) & \cdots & f_2^{(n-1)}(x) \\ \cdot & \cdot & \cdot & \cdot \\ f_n(x) & f_n'(x) & \cdots & f_n^{(n-1)}(x) \end{vmatrix}$$

by $W(f_1(x), f_2(x), \cdots, f_n(x))$. The equation (1) possesses in certain intervals the

Property W. *There exist* $n-1$ *integrals* $h_1(x)$, $h_2(x)$, $\cdots$, $h_{n-1}(x)$ *of the homogeneous equation* (1) *satisfying the* $n-1$ *inequalities*

$$h_1(x) > 0, \qquad W(h_1(x), h_2(x)) > 0, \quad \cdots, W(h_1(x), h_2(x), \cdots, h_{n-1}(x)) > 0 \tag{3}$$

throughout the open interval (a, b).

Determine $n-1$ particular integrals of (1) by such initial conditions that the $n-1$ Wronskians in question are $=1$ for $x = x_0$; then the inequalities (3) are satisfied in a sufficiently small interval surrounding the point x_0.

* Presented to the Society, October 27, 1923.

Equation (2) possesses the property W in the interval $(-\infty, +\infty)$ since the functions $1, x, x^2, \cdots, x^{n-2}$ are integrals of (2) and we have

$$1 > 0, \quad W(1, x) = 1! > 0, \quad W(1, x, x^2) = 1!\,2! > 0, \quad \cdots.$$

The equation

$$y'' + y = 0 \tag{4}$$

is of the second order. Thus there is only one inequality (3), which is satisfied by the integral $\sin(x - a)$ for $a < x < a + \pi$. No integral of the equation (4) is > 0 throughout an interval (a, b) where $b - a > \pi$. Therefore the property W belongs to (4) in any interval of length $\leqq \pi$ but in no interval of length $> \pi$. An equation $y' + \phi y = 0$ of the first order possesses the property W in any interval where ϕ is continuous.

We assume now (and also in §§ 3–5) *the existence of the property* W. Then we have the following theorems I, II, III:

THEOREM I. *If the function* $f(x)$, *supposed differentiable* n *times, vanishes at* $n + 1$ *points of the interval* (a, b), *then there exists an intermediate point* ξ *such that*

$$If(\xi) = f^{(n)}(\xi) + \phi_1(\xi) f^{(n-1)}(\xi) + \cdots + \phi_n(\xi) f(\xi) = 0.$$

Theorem I is a generalization of Rolle's theorem. Apply I to the equation (2) putting $n = 1$; then we have the ordinary theorem of Rolle.

THEOREM II. *There exists one and only one integral* $H(x)$ *of the homogeneous equation* (1) *assuming* n *given values at* n *given points of the interval* (a, b).

Theorem II asserts the possibility of a certain interpolation; in the case of equation (2) and if the n points are distinct, it asserts the possibility of solving the problem which is solved explicitly by Lagrange's formula.

THEOREM III. *Determine an integral* $H(x)$ *of the homogeneous equation* (1) *assuming the same values as* $f(x)$ *at* n *given points of* (a, b); *determine further an integral* $N(x)$ *of the non-homogeneous equation*

$$Ly \equiv \frac{d^n y}{dx^n} + \phi_1 \frac{d^{n-1} y}{dx^{n-1}} + \cdots + \phi_n y = 1, \tag{5}$$

that vanishes at the n *points in question. There exists a point* ξ *intermediate between these* n *points and an arbitrary point* x *of* (a, b) *such that*

$$f(x) = H(x) + N(x) If(\xi).$$

Take the equation (2) and n coincident points; in this particular case Theorem III gives Taylor's formula with Lagrange's remainder-term.

In all the theorems just stated coinciding points are admissible; this will be fully explained in § 1. § 2 enumerates some useful formulas concerning

Wronskians. The proofs of Theorems I, II, III and a few illustrative examples are given in §§ 3–5. In § 6 I try to justify the definition of property W by proving the converse of Theorem II. § 7 deals with a system of several functions of one variable and generalizes Rolle's theorem in another manner.

1. The function $f(x)$ is supposed differentiable n times. We say that $f(x)$ vanishes at k points coinciding with x_0 if

$$f(x_0) = f'(x_0) = \cdots = f^{(k-1)}(x_0) = 0. \tag{6}$$

This statement has a definite meaning if $k \leqq n+1$. If we say that $f(x)$ vanishes at k *and no more* points coinciding with x_0, then we add to (6) the condition

$$f^{(k)}(x_0) \neq 0,$$

and we suppose $k \leqq n$. If $f(x)$ vanishes at

$$m_1,\ m_2,\ m_3,\ \cdots,\ m_l \tag{7}$$

points coinciding with

$$x_1,\ x_2,\ x_3,\ \cdots,\ x_l \tag{8}$$

respectively where

$$a < x_1 < x_2 < x_3 < \cdots < x_l < b \tag{9}$$

then we say that $f(x)$ vanishes at $m_1 + m_2 + \cdots + m_l$ points of the interval (a, b). If we say that $f(x)$ and $\phi(x)$ assume the same values at n points of (a, b) we mean that $f(x) - \phi(x)$ vanishes at n points of (a, b). Theorem II may be stated more explicitly thus: if there are given l points (8) satisfying (9), l positive integers (7) such that $m_1 + m_2 + \cdots + m_l = n$, and finally l systems of values

$$\begin{array}{l} y_1,\ y_1',\ y_1'',\ \cdots,\ y_1^{(m_1-1)}, \\ y_2,\ y_2',\ y_2'',\ \cdots,\ y_2^{(m_2-1)}, \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ y_l,\ y_l',\ y_l'',\ \cdots,\ y_l^{(m_l-1)}, \end{array}$$

then there exists an integral $H(x)$ of (1) satisfying the conditions

$$\begin{array}{llll} H(x_1) = y_1, & H'(x_1) = y_1', & \cdots, & H^{(m_1-1)}(x_1) = y_1^{(m_1-1)}, \\ H(x_2) = y_2, & H'(x_2) = y_2', & \cdots, & H^{(m_2-1)}(x_2) = y_2^{(m_2-1)}, \\ \cdots\cdots & \cdots\cdots & \cdots\cdots & \cdots\cdots \\ H(x_l) = y_l, & H'(x_l) = y_l', & \cdots, & H^{(m_l-1)}(x_l) = y_l^{(m_l-1)}, \end{array} \tag{10}$$

and $H(x)$ is completely determined by these conditions.

A point ξ satisfying the inequality $x_1 < \xi < x_l$ is said to be *intermediate* between the points (8); so far we have supposed $l > 1$. If $l = 1$, that is, if all the points in question coincide with one point x_1, then $\xi = x_1$ is the only *intermediate* point.

I enumerate a few evident propositions used in what follows. The functions $u, u_1, \cdots, u_k$ are supposed continuous and differentiable as many times as required to calculate the expression

$$(11)\qquad \begin{aligned} & u_k \frac{d}{dx}\left(u_{k-1}\frac{d}{dx}\left(u_{k-2}\cdots\frac{d}{dx}\left(u_1\frac{d}{dx}(uf)\right)\cdots\right)\right) \\ & = u_k\frac{d}{dx}u_{k-1}\frac{d}{dx}u_{k-2}\cdots\frac{d}{dx}u_1\frac{d}{dx}uf = uu_1u_2\cdots u_k\frac{d^kf}{dx^k}+\cdots . \end{aligned}$$

1. *If $f(x)$ vanishes at k (and no more) points coinciding with x_0, then $f'(x)$ vanishes at $k-1$ (and no more) points coinciding with x_0.*

2. *If $f(x)$ vanishes at k points coinciding with x_0, then $u(x)f(x)$ vanishes at k (or more) points coinciding with x_0.*

3. *Suppose $f(x)$ vanishes at k and no more points coinciding with x_0, and suppose further that $u, u_1, u_2, \cdots, u_k$ are different from zero for $x = x_0$. Then the expression* (11) *is $\neq 0$ for $x = x_0$.*

4. *Suppose $f(x)$ vanishes at $k+1$ points. Then, between these $k+1$ points, there exists an intermediate point ξ at which the expression* (11) *vanishes.*

We obtain 4 by repeated application of Rolle's theorem and of Theorems 1, 2.

2. Let $c_1, c_2, \cdots, c_{n-1}$ be constants; then we have

$$(12)\qquad \begin{aligned} & W(f_1, f_2, \cdots, f_{n-1}, c_1f_1 + c_2f_2 + \cdots + c_{n-1}f_{n-1} + f_n) \\ & = W(f_1, f_2, \cdots, f_{n-1}, f_n). \end{aligned}$$

Suppose the function u is differentiable $n-1$ times; then we have

$$W(uf_1, uf_2, \cdots, uf_n) = u^n W(f_1, f_2, \cdots, f_n).$$

In particular put $u = 1/f_1$; we obtain

$$(13)\qquad W\left[\left(\frac{f_2}{f_1}\right)', \left(\frac{f_3}{f_1}\right)', \cdots, \left(\frac{f_n}{f_1}\right)'\right] = \frac{W(f_1, f_2, \cdots, f_n)}{f_1^n}.$$

For the derivative of a Wronskian we have

$$(14)\qquad W'(f_1, f_2, \cdots, f_n) = \left|f_k, f_k', \cdots, f_k^{(n-2)}, f_k^{(n)}\right|_{k=1, 2, \ldots, n}.$$

Using (14) we obtain by the usual formula for a minor of the adjoint determinant

$$(15)\qquad \begin{aligned} & W(f_1, \cdots, f_{n-2}, f_{n-1})W'(f_1, \cdots, f_{n-2}, f_n) \\ & \quad - W'(f_1, \cdots, f_{n-2}, f_{n-1})W(f_1, \cdots, f_{n-2}, f_n) \\ & \quad = W(f_1, \cdots, f_{n-2})W(f_1, \cdots, f_{n-2}, f_{n-1}, f_n), \end{aligned}$$

$$\begin{aligned} & \frac{d}{dx}\frac{W(f_1, \cdots, f_{n-2}, f_n)}{W(f_1, \cdots, f_{n-2}, f_{n-1})} \\ & \quad = \frac{W(f_1, \cdots, f_{n-2})W(f_1, \cdots, f_{n-2}, f_{n-1}, f_n)}{(W(f_1, \cdots, f_{n-2}, f_{n-1}))^2}. \end{aligned}$$

If $h_1(x), h_2(x), \cdots, h_n(x)$ are n linearly independent integrals of (1), we have

$$Ly \equiv \frac{W(h_1, h_2, \cdots, h_n, y)}{W(h_1, h_2, \cdots, h_n)}, \tag{16}$$

$$W(h_1, h_2, \cdots, h_n) = ce^{-\int \phi_1(x)dx}, \tag{17}$$

where c is a constant. Put

$$\begin{aligned} y &= Y_0, & W(h_1, y) &= Y_1, \\ W(h_1, h_2, y) &= Y_2, & \cdots, \quad W(h_1, h_2, \cdots, h_n, y) &= Y_n, \\ 1 &= W_0, & h_1 &= W_1, \\ W(h_1, h_2) &= W_2, & \cdots, \quad W(h_1, h_2, \cdots, h_n) &= W_n. \end{aligned}$$

Applying (15) we obtain

$$\frac{d}{dx}\frac{Y_0}{W_1} = \frac{W_0 Y_1}{W_1^2}, \qquad \frac{d}{dx}\frac{Y_1}{W_2} = \frac{W_1 Y_2}{W_2^2}, \qquad \cdots, \qquad \frac{d}{dx}\frac{Y_{n-1}}{W_n} = \frac{W_{n-1} Y_n}{W_n^2}.$$

Combining these formulas with (16) we find finally

$$Ly \equiv \frac{W_n}{W_{n-1}} \frac{d}{dx} \frac{W_{n-1}^2}{W_{n-2} W_n} \cdots \frac{d}{dx} \frac{W_2^2}{W_1 W_3} \frac{d}{dx} \frac{W_1^2}{W_0 W_2} \frac{d}{dx} \frac{y}{W_1}. \tag{18}$$

(18) is the usual decomposition of a linear homogeneous differential expression into a "product" of differential expressions of the first order. The proof here given is perhaps more direct than the usual.*

3. The property W implies the non-vanishing of $W_1, W_2, \cdots, W_{n-1}$ appearing in formula (18). W_n vanishes nowhere, by virtue of (17). Applying Theorem 4 of § 1 to the decomposition (18) we get at once Theorem I.

Apply now Theorem I to $f(x) - H(x)$ instead of $f(x)$, where $H(x)$ is an integral of the homogeneous equation (1). Then $LH(x) = 0$ and we obtain the corollary

THEOREM I*. *If the function $f(x)$, supposed differentiable n times, assumes at $n+1$ points of the interval (a, b) the same values as an integral of the homogeneous equation* (1), *there exists an intermediate point ξ such that*

$$Lf(\xi) = f^{(n)}(\xi) + \phi_1(\xi) f^{(n-1)}(\xi) + \cdots + \phi_n(\xi) f(\xi) = 0.$$

Theorem I* was incidentally remarked by H. Poincaré† in the case of an equation of the second order where the property W is particularly simple, including only one inequality. Poincaré's proof is very different from the proof given here and more complicated. I obtained the results presented in

* Cf. L. Schlesinger, *Handbuch der Theorie der linearen Differentialgleichungen*, vol. I, p. 52, formula (14).

† L'Intermédiaire des Mathématiciens, vol. 1 (1894), pp. 141–144 and also pp. 69, 127, 172, 216.

this paper in clearing up and working out the original remark made by Poincaré.

Another corollary to Theorem I is the following

THEOREM I$**$. *No integral $N(x)$ of the non-homogeneous equation* (5) *vanishes at more than n points of the interval (a, b).*

Suppose, if possible, that $N(x)$ vanishes at $n+1$ points; then we have by Theorem I

$$LN(\xi) = 0$$

at an intermediate point ξ; now, actually, $LN(\xi) = 1$ and thus Theorem I$**$ is proved.

4. The general integral of (1) has the form

$$H(x) = c_1 h_1(x) + c_2 h_2(x) + \cdots + c_n h_n(x),$$

where $c_1, c_2, \cdots, c_n$ are constants. The conditions (10) form a system of n linear non-homogeneous equations with the aid of which we have to determine $c_1, c_2, \cdots, c_n$. We have to prove that this system is neither incompatible nor indeterminate. Thus we have to prove that the only solution of the *homogeneous* system obtained from (10) by putting

$$y_1 = y_1' = y_2'' = \cdots = y_i^{(m_i-1)} = 0 \tag{19}$$

is the identically vanishing solution.

Suppose, if possible, that there is a solution of the homogeneous system (10) (19) which does not vanish identically; denote it by

$$H(x) = \gamma_1 h_1(x) + \gamma_2 h_2(x) + \cdots + \gamma_k h_k(x) + \cdots + \gamma_n h_n(x)$$

where

$$\gamma_k \neq 0, \qquad \gamma_{k+1} = \gamma_{k+2} = \cdots = \gamma_n = 0.$$

Then $h_k(x)$ assumes at $n \geqq (k-1)+1$ points the same values as an integral of the equation

$$\frac{W(h_1, h_2, \cdots, h_{k-1}, y)}{W(h_1, h_2, \cdots, h_{k-1})} = 0$$

of order $k-1$. Then there exists, by virtue of I$*$, an intermediate point ξ at which

$$W(h_1(\xi), h_2(\xi), \cdots, h_{k-1}(\xi), h_k(\xi)) = 0.$$

This last equation is excluded if we assume the existence of the property W and thus Theorem II is proved.

Example. Assume that the algebraic equation

$$x^n + c_1 x^{n-1} + \cdots + c_n = 0$$

has n different real roots $a_1, a_2, \cdots, a_n$, where

$$a_1 < a_2 < a_3 < \cdots < a_n. \tag{20}$$

Then the linear homogeneous equation with constant coefficients

$$y^{(n)} + c_1 y^{(n-1)} + c_2 y^{(n-2)} + \cdots + c_n y = 0$$

has the integrals

$$e^{a_1 x},\ e^{a_2 x},\ \cdots,\ e^{a_n x}, \tag{21}$$

and possesses the property W in the interval $(-\infty, +\infty)$ since the Wronskians

$$e^{a_1 x},\ W(e^{a_1 x}, e^{a_2 x}),\ \cdots,\ W(e^{a_1 x}, e^{a_2 x}, \cdots, e^{a_n x})$$

divided by

$$e^{a_1 x},\ e^{(a_1+a_2)x},\ \cdots,\ e^{(a_1+a_2+\cdots+a_n)x}$$

respectively reduce to positive constants which may be calculated by Vandermonde's formula. Therefore we can interpolate by a suitably chosen linear combination of the n integrals (21) n arbitrary values at the n points $x_1, x_2, \cdots, x_n$ where

$$x_1 < x_2 < x_3 < \cdots < x_n. \tag{22}$$

This fact is equivalent to the statement that

$$\begin{vmatrix} e^{a_1 x_1} & e^{a_1 x_2} & \cdots & e^{a_1 x_n} \\ e^{a_2 x_1} & e^{a_2 x_2} & \cdots & e^{a_2 x_n} \\ \cdot & \cdot & \cdot & \cdot \\ e^{a_n x_1} & e^{a_n x_2} & \cdots & e^{a_n x_n} \end{vmatrix} \neq 0. \tag{23}$$

Thus we have proved (23) as a particular case of Theorem II. It may be shown directly in a slightly different manner that the determinant (23) is > 0 if (20) (22) are satisfied. I add that the property W belongs to any equation whatever with constant coefficients in the interval $(-\infty, +\infty)$ if the roots of the corresponding algebraic equation are all real, *whether distinct or not*. We get a simple proof of this from (17).

Any integral $N(x)$ of the non-homogeneous equation (5) is of the form

$$N(x) = N_0(x) + H(x).$$

Here $N_0(x)$ is a particular integral of the non-homogeneous equation (5) and $H(x)$ is the general integral of the homogeneous equation (1); $H(x)$ contains n arbitrary constants which we can adapt to any conditions of the form (10), by virtue of Theorem II. Thus we obtain the corollary

THEOREM II$*$. *There exists one and only one integral $N(x)$ of the non-homogeneous equation* (5) *assuming n given values at n given points of the interval* (a, b).

5. I now prove Theorem III. The possibility of determining $H(x)$ according to the enunciated conditions follows from Theorem II and the possibility

of determining $N(x)$ follows from II$*$. Let x_0 be a point of (a, b) but not one of the n points mentioned at which $f(x) - H(x)$ and $N(x)$ vanish. Then we have, by Theorem I$**$, $N(x_0) \neq 0$. Therefore it is possible to find a constant C satisfying the equation

$$f(x_0) = H(x_0) + CN(x_0). \tag{24}$$

Thus the function

$$f(x) - H(x) - CN(x)$$

vanishes at $n + 1$ *points* of (a, b). There exists, by Theorem I, an intermediate point ξ at which

$$L(f(\xi) - H(\xi) - CN(\xi)) = Lf(\xi) - 0 - C \cdot 1 = 0,$$

that is,

$$C = Lf(\xi).$$

Substituting this value in (24) we have Theorem III.*

I will now give some concrete cases of Theorem III with immediate applications.

Example 1. Let x_1 be the point at which we interpolate the differentiable function $f(x)$ by an integral of the equation

$$y' + y = 0.$$

Adopting the notation of Theorem III we have

$$H(x) = f(x_1)e^{x_1 - x}, \qquad N(x) = 1 - e^{x_1 - x},$$

and consequently for $x > x_1$

$$f(x) = f(x_1)e^{x_1 - x} + (1 - e^{x_1 - x})(f(\xi) + f'(\xi)) \tag{25}$$

where $x_1 < \xi < x$. From (25) we get at once the proposition †

Suppose $f(x)$ *differentiable for* $x > 0$; *then*

$$\lim_{x \to +\infty} (f(x) + f'(x)) = c$$

implies $\lim f(x) = c$.

We may assume $c = 0$ (and apply the proof to $f(x) - c$ instead of to $f(x)$). Choose first x_1 sufficiently great so that $|f(x) + f'(x)| < \epsilon/2$ for $x > x_1$; then choose x_2, $x_2 > x_1$, such that $|f(x_1)|e^{x_1 - x_2} < \epsilon/2$. Finally we get by (25) for $x > x_2$

$$|f(x)| < \frac{\epsilon}{2} + \frac{\epsilon}{2}.$$

* The particular case corresponding to (2) was given by Ch. Hermite for points coinciding in an arbitrary manner. See *Œuvres*, vol. III, pp. 432–443, and T. J. Stieltjes, *Œuvres*, vol. I, pp. 47–60. For a more general point of view see G. D. Birkhoff, these Transactions, vol. 7 (1906), pp. 107–136.

† G. H. Hardy, Quarterly Journal of Mathematics, vol. 35 (1903), pp. 31–33. The proof given by O. Perron, Mathematische Zeitschrift, vol. 6 (1920), p. 159, supposes *more* than the derivability of $f(x)$.

By the same argument we may prove an analogous result for any equation with constant coefficients if the roots of the corresponding algebraic equation are real and negative.*

Example 2. In dealing with the equation

$$y'' + y = 0 \tag{4}$$

we may choose for (a, b) any interval of length π. The two points at which we interpolate may be different or not. Interpolating at two distinct points x_1, x_2, where

$$x_1 < x_2 < x_1 + \pi, \tag{26}$$

we find the formula

$$f(x) = \frac{f(x_1)\sin(x_2 - x) + f(x_2)\sin(x - x_1)}{\sin(x_2 - x_1)} + \left(1 - \frac{\cos\left(x - \frac{x_1 + x_2}{2}\right)}{\cos\frac{x_2 - x_1}{2}}\right)(f''(\xi) + f(\xi)) \tag{27}$$

valid for $x_2 - \pi < x < x_1 + \pi$; ξ is an interior point of the least interval containing x_1, x_2, x. Interpolating at two points coinciding with x_1 we find

$$f(x) = f(x_1)\cos(x - x_1) + f'(x_1)\sin(x - x_1) + (1 - \cos(x - x_1))(f''(\xi) + f(\xi)) \tag{28}$$

valid for $x_1 - \pi < x < x_1 + \pi$; ξ lies between x_1 and x. We see that restriction (26) is essential for the validity of (27), $x_2 = x_1 + \pi$ being generally inadmissible; this point will be cleared up by Theorem IV.

In order to give an application suppose $f(x)$ is an integral of the equation

$$y'' + \phi(x)y = 0, \tag{29}$$

where

$$\phi(x) > 1; \tag{30}$$

suppose further that $f(x_1) > 0$. We have by (29)

$$f''(\xi) + f(\xi) = -(\phi(\xi) - 1)f(\xi);$$

thus we may put (28) in the form

$$f(x) = A\sin(x - a) - (1 - \cos(x - x_1))(\phi(\xi) - 1)f(\xi) \tag{31}$$

* O. Perron, Mathematische Zeitschrift, vol. 6 (1920), pp. 161–163.

where

$$(32) \qquad A \sin (x_1 - a) = f(x_1) > 0, \qquad A \cos (x_1 - a) = f'(x_1).$$

The situation shown by (30) (31) (32) is this: the curve $A \sin (x - a)$ meets the curve $f(x)$ at the point x_1, see (32); the curve $A \sin (x - a)$ runs on the upper side of the curve $f(x)$ *as far as* $f(x)$ runs on the upper side of the x-axis, see (31) (30); but the curve $A \sin (x - a)$ cuts the x-axis at two points between which x_1 lies and whose distance apart is π; hence $f(x)$ reaches the x-axis necessarily *before* $A \sin (x - a)$ *cuts it* and we have the theorem

The distance between two consecutive zeros of any integral of the equation (29) *satisfying the condition* (30) *is* $< \pi$.

This is a particular case of a well known classical theorem given by Sturm* that can be proved generally by the argument I have used in the foregoing concrete case.

Any integral $f(x)$ of the equation (29) has the following property: the inequality

$$(33) \qquad f(x) > 0$$

implies

$$(34) \qquad f''(x) + f(x) < 0.$$

Consider the curve of which the equation is

$$(35) \qquad r = \frac{1}{f(\vartheta)}$$

in a system of polar coördinates r, ϑ with origin O. The inequality (34) expresses that the curve (35) appears from O to be *convex*, as we may see, e.g., by (27). The result just proved may be formulated geometrically as follows:

A curve appearing from a given point O to be everywhere convex subtends an angle $< \pi$ *at O*.

H. Poincaré was led to the remark above mentioned (§ 3) by constructing an analytical demonstration of this geometrically evident fact.

6. The converse of Theorem II is also true: if the interpolation of any n values by an integral of (1) is possible at any n points of an interval, then the equation (1) possesses the property W in that interval. The possibility of interpolation, that is, the *existence* of a solution of the non-homogeneous system (10) with arbitrary $y_1, y_1', \cdots, y_1^{(m_1-1)}$, is equivalent to the *non-existence* of a non-vanishing solution of the corresponding homogeneous system; we used this equivalence in § 4. Thus the most simple form of the converse to Theorem II is the following:

* Cf. Maxime Bôcher, *Méthodes de Sturm*, Paris, 1917, p. 52.

THEOREM IV. *If, except the identically vanishing integral, there exists no integral of equation* (1) *vanishing at n points of the interval* $a \leqq x < b$, *then there exist* $n-1$ *integrals* $h_1(x), h_2(x), \cdots, h_{n-1}(x)$ *such that the Wronskians*

(36) $\quad h_1(x),\ W(h_1(x), h_2(x)),\ \cdots,\ W(h_1(x), h_2(x), \cdots, h_{n-1}(x))$

vanish nowhere in the interval $a < x < b$.

The signs of the Wronskians are of no importance provided that they are constant. If the Wronskians (36) do not vanish we can make them positive by substituting $-h_\nu(x)$ for $h_\nu(x)$ for certain values of ν.

Determine n integrals $h_1(x), h_2(x), \cdots, h_n(x)$ satisfying the initial conditions

$$h_1(a) = h_1'(a) = h_1''(a) = \cdots = h_1^{(n-3)}(a) = h_1^{(n-2)}(a) = 0, \quad h_1^{(n-1)}(a) = 1,$$

$$\text{(37)} \quad \begin{aligned} &h_2(a) = h_2'(a) = h_2''(a) = \cdots = h_2^{(n-3)}(a) = 0,\ h_2^{(n-2)}(a) = 1, \\ &h_{n-2}(a) = h_{n-2}'(a) = 0, \quad h_{n-2}''(a) = 1, \\ &h_{n-1}(a) = 0, \quad h_{n-1}'(a) = 1, \\ &h_n(a) = 1. \end{aligned}$$

These integrals are linearly independent because their Wronskian $= \pm 1$ for $x = a$; $h_1(x)$ vanishes at $n-1$ points coinciding with a; by hypothesis there exists no nth point in the interval $a \leqq x < b$ at which $h_1(x)$ vanishes; therefore $h_1(x) \neq 0$ for $a < x < b$. Adopt the notation

$$W(h_1, h_2, \cdots, h_k) = W_k$$

that we used in § 2 and suppose that we have already proved

$$\text{(38)} \qquad W_1 \neq 0, \quad W_2 \neq 0, \quad \cdots, \quad W_k \neq 0$$

for $a < x < b$, $k < n$.

Let $c_1, c_2, \cdots, c_k$ be constants and put

$$c_1 h_1(x) + c_2 h_2(x) + \cdots + c_k h_k(x) + h_{k+1}(x) = h(x).$$

$h(x)$ does not vanish identically, and therefore it vanishes at no more than $n-1$ points contained in the interval $a \leqq x < b$. Now $h(x)$ vanishes at $n-k-1$ points coinciding with a by virtue of the conditions (37). Determine the constants $c_1, c_2, \cdots, c_k$ in such a manner that $h(x)$ vanishes at k points coinciding with x_0 where $a < x_0 < b$; this is possible because the determinant of the k equations in question

$$W(h_1(x_0), h_2(x_0), \cdots, h_k(x_0)) \neq 0,$$

by the hypothesis of complete induction, that is, by (38). Beyond the

$$(n-k-1) + k = n-1$$

zeros already mentioned, $h(x)$ has no other zeros for $a \leqq x < b$. In particular $h(x)$ *vanishes at no more than k points coinciding with* x_0. Consequently, by virtue of hypothesis (38) and of the Theorem 3 of § 1,

$$(39)\qquad \frac{W_k^2}{W_{k-1}}\frac{d}{dx}\frac{W_{k-1}^2}{W_{k-2}W_k}\cdots\frac{d}{dx}\frac{W_2^2}{W_1W_3}\frac{d}{dx}\frac{W_1^2}{W_0W_2}\frac{d}{dx}\frac{h}{W_1} = W(h_1, h_2, \cdots, h_k, h) = W(h_1, h_2, \cdots, h_k, h_{k+1}) = W_{k+1}$$

does not vanish for $x = x_0$; we obtain (39) from (15) (12) in the same manner as we obtained (18). The point x_0 was arbitrarily chosen in the interval $a < x < b$ and thus Theorem IV is proved.

7. Another demonstration of the Theorems I, II, III essentially different from that given in §§ 1–5 is based on the following theorem,

THEOREM V. *Suppose all the functions* $h_1(x), h_2(x), \cdots, h_n(x), f(x)$ *possess derivatives of order n throughout the interval* $a < x < b$ *and satisfy the following n inequalities:*

$$h_1(x) > 0, \qquad W(h_1(x), h_2(x)) > 0, \qquad \cdots, \qquad W(h_1(x), h_2(x), \cdots, h_n(x)) > 0.$$

Suppose also that the l points $x_1, x_2, \cdots, x_l$ *and the corresponding positive integers* $m_1, m_2, \cdots, m_l$ *satisfy the conditions*

$$a < x_1 < x_2 < \cdots < x_l < b, \qquad m_1 + m_2 + \cdots + m_l = n + 1.$$

Consider the determinant of order $n + 1$

$$(40)\qquad \left|\begin{array}{c} h_k(x_1), h_k'(x_1), \cdots, h_k^{(m_1-1)}(x_1), h_k(x_2), h_k'(x_2), \\ \cdots, h_k^{(m_2-1)}(x_2), \cdots, h_k(x_l), \cdots, h_k^{(m_l-1)}(x_l) \end{array}\right|,$$

the n first rows being obtained by putting successively $k = 1, 2, 3, \cdots, n$ *and the last row by substituting* $f(x)$ *for* $h_k(x)$.

There exists an intermediate point ξ, $x_1 < \xi < x_l$, *such that the value assumed by the Wronskian at the point* ξ, *that is,*

$$W(h_1(\xi), h_2(\xi), \cdots, h_n(\xi), f(\xi)),$$

is $>$, $=$ *or* < 0, *according as the determinant* (40) *is* $>$, $=$ *or* < 0.

Theorem V may be compared to a well known determinantal theorem of H. A. Schwarz.* Theorem V may be proved by complete induction; its validity for the $(n-1)+1$ functions

* H. A. Schwarz, *Abhandlungen*, vol. II, pp. 296–302; T. J. Stieltjes, *Œuvres*, vol. II, pp. 110–123. The determinant (40) has been considered by H. M. Morse, G. A. Pfeiffer, and G. M. Green from a different point of view; see Bulletin of the American Mathematical Society, vol. 23 (1916), pp. 114–122.

$$\left(\frac{h_2(x)}{h_1(x)}\right)', \left(\frac{h_3(x)}{h_1(x)}\right)', \cdots, \left(\frac{h_n(x)}{h_1(x)}\right)', \left(\frac{f(x)}{h_1(x)}\right)'$$

implies its validity for the $n+1$ functions $h_1(x), h_2(x), \cdots, h_n(x), f(x)$, as may be shown by (13) and by Rolle's theorem. I had originally based my demonstration of Theorems I, II, III on Theorem V. I was led to the treatment of the subject I finally adopted by a kind remark made by Professor H. Weyl.

EIDG. TECHNISCHE HOCHSCHULE,
ZURICH, SWITZERLAND.

ERRATA, VOLUME 24

J. F. RITT, *On algebraic functions which can be expressed in terms of radicals*. Page 21, lines 30 and 33, for "n^2" read "n".

MATHEMATICAL SCIENCES LIBRARY
SIDNEY SMITH HALL
UNIVERSITY OF TORONTO

A.M.S.

3557-9 Year 1922

Richard Courant

ST. MIKE'S

Postal Code

Date

GretagMacbeth™ ColorChecker Color Rendition Chart